January 2016

Defining the regulatory determinants in substrate catalysis by biochemical, biophysical, and kinetic studies for the development of specific small-molecule inhibitors of ubiquitin specific proteases 7 and 17

Nicole M. Hjortland
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By NICOLE M HJORTLAND

Entitled

DEFINING THE REGULATORY DETERMINANTS IN SUBSTRATE CATALYSIS BY BIOCHEMICAL, BIOPHYSICAL, AND KINETIC STUDIES FOR THE DEVELOPMENT OF SPECIFIC SMALL-MOLECULE INHIBITORS OF UBIQUITIN SPECIFIC PROTEASES 7 AND 17

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

ANDREW MESECAR
Chair

HUMAIRA GOWHER

JEFFREY BOLIN

TIMOTHY RATLIFF

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): DR. ANDREW MESECAR

Approved by: CHRISTINE HRYCYNA 9/14/2016

Head of the Departmental Graduate Program Date

DEFINING THE REGULATORY DETERMINANTS IN SUBSTRATE CATALYSIS BY BIOCHEMICAL, BIOPHYSICAL, AND KINETIC STUDIES FOR THE DEVELOPMENT OF SPECIFIC SMALL–MOLECULE INHIBITORS OF UBIQUITIN SPECIFIC PROTEASE 7 AND 17

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Nicole M. Hjortland

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2016

Purdue University

West Lafayette, Indiana

For my brilliant and loving husband, Andy,
and my strong and supportive mom, Judy.

ACKNOWLEDGMENTS

My journey through graduate school was not travelled alone by any means. In fact, it took a whole team of people supporting me, challenging me, and picking me up when I was down to get to this point, and proper gratitude must be expressed.

First and foremost, I would like to my advisor Dr. Andrew Mesecar for taking a chance on me. Throughout my time as a member of your lab, it seemed that you saw my true potential, even when I could not. Your motivation to keep going allowed me to push the scientific frontier forward. Thank you to my committee members Dr. Jeffrey Bolin, Dr. Chittaranjan Das, Dr. Timothy Ratliff, and Dr. Humaira Gowher for challenging me to think outside of my comfort zone. It was through your guidance and our scientific discussions that the depth of my scientific knowledge grew.

Next I would like to thank those I worked closely with over the years, and who helped to move the project along. This includes Dr. Antonella Pepe, Dr. Larisa Avramova, and Dr. Scott Crist. Dr. Pepe was always a friendly face as she single-handedly synthesized all the compounds utilized in Chapter 6. Dr. Avramova helped me to the design and execute the performed high throughput screens in Chapter 5 and Chapter 6. She was incredibly patient and thorough in teaching me the many aspects of this important methodology. Dr. Scott Crist and I attempted to evaluate the USP7 compounds in cell culture. While these experiments proved to be very difficult, throughout the process I could always count on Dr. Scott Crist for his renowned enthusiasm.

While slightly unconventional, I would like to thank my high school chemistry teacher, Mrs. Cindy Carter. Of all the teachers and mentors I have had along the way, she was the first to spark my interest in science and she instilled in me a passion for science that I could not ignore.

I would like to acknowledge the current and past members of the Mesecar Lab Dr. Yahira Baez–Santos, Dr. Katherine Jensen, Dr. Sergey Savinov, Dr. Soma Mukhopadhyay, Dr. Laura Kingsley, Dr. Yafang Chen, Dr. Sakshi Tomar, Julia Luciano–Chadee, Qing Zhou, Yu–Chen Yen, Aya Ryuzoji, Dr. Sarah St. John, Dr. Renata Everett, Brandon Anson, Jozlyn Clasman, Corey Moore, Thays Carvalho, and Dr. Kristina Kesely. Thank you so much for being influential colleagues, I am so glad our paths crossed within the Mesecar Lab. Thank you to my undergraduate students, Anuj Patel and Julia Weeder, for teaching me how to mentor and making the lab an entertaining place to work. A special thank you to Dr. Yahira Baez–Santos for helping me get my start and being a wonderful source of support and friendship. Thank you to the great friends I have made in the lab, Yafang Chen, Sakshi Tomar, and Qing Zhou for their kind words, wisdom, and friendship. Thank you to Julia Luciano–Chadee for being awesome and reminding me to live life without regrets. A huge thank you to my partner in crime Dr. Sarah St. John without whom I doubt I would have finished my degree. You kept me going within the lab by making me smile and getting into shenanigans with me.

My friends outside of my lab here at Purdue are some of the finest people I have ever met including Amy Funk, Brad Jakubison, Stephanie Pratt, Evan Pratt, Dan Ysselstein, and Tiffany Young. Thank you so much for our adventures, our laughter, and our celebrations. I am so grateful to have had each of you as my friend along the way.

Finally, I could not have achieved this accomplishment without the unconditional love and support of my family. I was fortunate enough to have numerous Aunts and Uncles cheering me on as I traversed this degree. My late grandmother, Jeanette Davis, always reminded to have fun in life and enjoy the little things. My grandparents, Marylin and Gordon Ladwig, have been my pillars of strength throughout my life. My late father, Mark Davis, seemed to know my true potential early on in my life. He taught me to stand up for myself and encouraged me to fight for my dreams, to never give up. Jacob Davis has been an extremely supportive and loving

brother, always reminding me to do my best and try my hardest at everything I do in life. A special thank you to my mother, Judy Davis. I grew up watching her fight for her dreams and push past any barriers that stood in her way. It was through watching these actions as a young child that she, unknowingly, laid the groundwork early in my life for me to become a strong, independent, and determined woman. Finally, thank you to my husband, Andrew Hjortland. You are a patient, kind, and supportive husband. You held me up when I would stumble and have never stopped fighting by my side. I am truly blessed to be apart of such a loving, wonderful, and supportive family. Thank you all so much.

TABLE OF CONTENTS

Page

Page

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

ABBREVIATIONS

A_{280}	Absorbance at 280nm
A_{600}	Absorbance at 600nm
AICc	Akaike Information Criterion for small sample size
AcNPV	*Autographa californica* nuclear polyhedrosis virus
ATP	Adenosine triphosphate
AUC	Analytical Ultracentrifugation
BSA	Bovine Serum Albumin
C	Celsius
cm	Centimeter
CV	Column volume
d	Diameter
Da	Dalton
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
DLS	Dynamic Light Scattering
EDTA	Ethyenediaminetetraacetic Acid
FBS	Fetal Bovine Serum
FPLC	Fast Pressure Liquid Chromatography
g	Gram
GMPS	guanosine 5–monophosphate synthase
H1–5	USP7/HAUSP Ubiquitin–like Domains 1–5 (residues 560–1102)
H1–3	USP7/HAUSP Ubiquitin–like Domains 1–3 (residues 560–888)
H1–2	USP7/HAUSP Ubiquitin–like Domains 1–2 (residues 560–775)

H4–5	USP7/HAUSP Ubiquitin–like Domains 4–5 (residues 890–1102)
HEPES	4–(2–hydroxyethyl)–1–piperazineethanesulfonic acid
IC_{50}	Concentration required to inhibit enzyme activity by 50%
IPTG	Isopropyl β–D–1–thiogalactopyranoside
ITC	Isothermal titration calorimetry
LB Media	Luria–Bertina Media
MW	Molecular Weight
MWCO	Molecular Weight Cut Off
mAU	Milli–Absorbance Units
mM	millimolar
NaCl	Sodium Chloride
nM	Nanomolar
nm	Nanometer
OTG	Octylthioglucoside
PDB	Protein Data Bank
R_H	Hydrodynamic Radius
rpm	Revolutions Per Minute
SAR	Structure Activity Relationship
SDS	Sodium Dodecyl Sulfate
SDS–PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEC	Size–Exclusion Chromatography
TBS	Tris-Buffered Saline
TBST	Tris-Buffered Saline with 1% Tween20
TCEP	Tris(2–carboxyethyl)phosphine
TRAF	TNF receptor associated factor
TRAF–$USP7_{CD}$	Truncation of Ubiquitin Specific Protease 7 (residues 1–560)
Ub–AMC	Ubiquitin 7–amido–4–methylcoumarin
Ub–Rho110	Ubiquitin Rhodamine 110
Ubl	Ubiquitin Like Domains

USP	Ubiquitin Specific Protease
USP7	Ubiquitin Specific Protease 7
$USP7_{CD}$	Catalytic Domain Truncation (residues 208–560) Ubiquitin Specific Protease 7
$USP7_{FL}$	Full Length (residues 1–1102) Ubiquitin Specific Protease 7
$USP7_{CD}$–H1–5	Truncated $USP7_{CD}$–H1–5 Ubiquitin Specific Protease 7 (residues 208–1102)
$USP7_{CD}$–H4–5	Truncated $USP7_{CD}$–H4–5 Ubiquitin Specific Proteases (residues 208-1102 Δ567–879)
USP17	Ubiquitin Specific Protease 17
UV	Ultraviolet
V	Volts
μM	Micromolar

ABSTRACT

Hjortland, Nicole M. Ph.D., Purdue University, December 2016. Defining the Regulatory Determinants in Substrate Catalysis by Biochemical, Biophysical, and Kinetic studies for the Development of Specific Small–Molecule Inhibitors of Ubiquitin Specific Protease 7 and 17. Major Professor: Andrew D. Mesecar.

Ubiquitination is an important post–translational modification involved in maintaining cellular homeostasis by regulating many delicate cellular processes, including the cell–cycle, membrane protein trafficking, endocytosis and apoptosis. Ubiquitin Specific Proteases (USPs) remove ubiquitin modifications from protein substrates to reverse the signal imposed by the ubiquitination. Perturbations in the expression levels of USPs has been implicated in many types of cancers where patients show significant elevation in cellular levels of specific USPs. This suggests that targeting specific upregulated members of the USP family in specific disease states would be ideal for the development of personalized anti–cancer therapeutics.

Each member of the USP family is composed of a conserved catalytic domain flanked by a variety of other domains that make the function of each USP unique. These flanking domains are important for substrate recognition, regulation, and function. The conserved catalytic domains of the USPs share high structural similarities, and within two regions of the catalytic domain they also share high sequence similarity. These regions are referred to as the Cys and His boxes, and they contain the catalytic triad composed of a cysteine, histidine, and aspartate residues. These structural and sequence similarities make the task of inhibiting specific members of the USP family substantially more difficult. Therefore, it is essential that we thoroughly characterize not only the enzymatic activity of the catalytic domain of the specific USP of interest; but also the structure and function of the flanking domains that

make the USP unique. Along these lines, this dissertation describes the enzymatic and biochemical characterization of two USP family members, USP17 and USP7.

USP17 is a cytokine–inducible USP that is required for cell–cycle progression through the G_1–S and G_2–M checkpoints. No structural information specific to USP17 is available, but cellular studies have determined that USP17 contains the conserved catalytic domain at the N–terminus and two hyaluronan binding motifs at the C–terminus. Cellular characterization of USP17 has led to the identification of three substrates that USP17 deubiquitinates: the Ras Converting Enzyme 1 (RCE1), the Cell Division Cycle 25A (CDC25A) phosphatase, and a component of the Sin3 co–repressor complex SDS3. Under normal cellular conditions, USP17 is only expressed during the cell–cycle checkpoints and if USP17 is expressed at other times, cellular proliferation ceases. However, the upregulation of USP17 has been implicated in both breast and prostate cancer, and it is thought that the persistent overexpression of USP17 applies continuous stress on the cell–cycle through perturbation of the Ras and CDC25A signaling pathways. As such, USP17 is an important target for inhibition leading to the development of anti–cancer therapeutics.

Here the novel expression, refolding, and purification methodology for USP17 is reported. USP17 is highly insoluble when expressed in *E. coli* expression systems, and, as such, no methodology for the recombinant expression and purification had previously been reported. Numerous USPs have successfully been expressed from baculovirus expression systems, including the well characterized USP7. USP17 was successfully expressed and purified from the baculovirus expression system, however due to high levels of degradation, the yield of intact USP17 was low. Therefore, a suitable buffer was identified and optimized for the refolding of active USP17 expressed in *E. coli* inclusion bodies. After refolding, active intact USP17 was purified using two chromotographic steps to $\geq$ 98% purity. The ability to purify milligram quantities of USP17 allowed for the enzymatic and biochemical characterization of USP17 including the determination of the steady–state kinetic parameters, defining the oligomeric state, and substrate specificity. Finally, in collaboration with the

Biomolecular Screening and Drug Discovery Facility in Bindley Biosciences Center, the recombinatly expressed USP17 underwent and initial round of high–throughput screening for the identification of potential small–molecule inhibitors.

Ubiquitin Specific Protease 7 (USP7/HAUSP) is to date the most well–studied human deubiqutinating enzyme of the USP class that regulates many high profile signaling pathways in the cell, including regulation of the tumor suppressor p53/MDM2 pathway. When the cell is under normal conditions, USP7 primarily deubiquitinates MDM2. However, when the cell experiences genotoxic stress the focus of USP7s deubiquitination shifts from MDM2 to p53. Furthermore, under genotoxic stress, guanosine 5–monophosphate synthase (GMPS) binds to and allosterically activates USP7 to enhance the rate of deubiquitination and thereby stabilize p53. Disruption of the p53 signaling pathway is responsible for roughly 50% of human cancers due to mutation or inactivation of p53. USP7 has been identified as being upregulated in many forms of cancer, essentially silencing p53 through the inappropriate stabilization of MDM2. This is only one of the pathways that identify USP7 as a promising target for inhibition and development of anti–cancer therapeutics.

While no full length USP7 structure has been solved, there is significant structural data of the individual domains of USP7 that have aided in our understanding of USP7's enzymatic activity and protein–protein interactions. The catalytic domain of USP7 has significantly lower activity than the full length enzyme and requires intramolecular activation by the C–terminal ubiquitin–like domains (HUBL) 4 and 5. Previous reports have suggested that HUBL domains 1–3 are not involved in the intramolecular activation of USP7; but rather act as a scaffold for the binding of the allosteric activator GMPS for enhanced activation of USP7. Utilizing kinetic and biochemical techniques, the importance of HUBL domains 1–3 in the intramolecular activation of USP7 was elucidated, and the results are striking.

The role of the HUBL domains 1–3 in the intramolecular activation of USP7 was evaluated using nine different USP7 truncations: $USP7_{FL}$, $USP7_{CD}$, $USP7_{CD}$–H1–5, $USP7_{CD}$–H4–5, TRAF–$USP7_{CD}$, H1–5, H4–5, H1–2, and H1–3. With the exception

of $USP7_{FL}$ which was prepared from the baculovirus expression system, each truncation was expressed from *E. coli* and purified using three sequential chromotographic steps. All truncations were purified in milligram quantities with $\geq 98\%$ purity. The steady–state kinetic parameters were determined for the active truncations ($USP7_{FL}$, $USP7_{CD}$, $USP7_{CD}$–H4–5, and TRAF–$USP7_{CD}$) with the fluorogenic substrate Ub–Rho110. The values obtained validated the requirement of HUBL domains 4 and 5 to achieve full activity of USP7. Interestingly, the parameters associated with the TRAF–$USP7_{CD}$ truncation suggest an interaction between Ub–Rho110 and the TRAF domain that has not been previously reported. Fixed yet variable kinetic studies defined striking differences between the nonessential activation parameters, $\alpha K_{A,app}$, $\alpha K_{m,app}$, α–, and β–, of $USP7_{CD}$ *in trans* by either H1–5 and H4–5. Upon further biochemical evaluation, it was determined that H1–3 act as a rheostat to regulate the level of activation H4–5 can apply to USP7.

A promising small molecule inhibitor of USP7 was identified for $USP7_{FL}$ by HTS. A structure activity relationship in regards to the halogen substiutent within the identified scaffold was developed in collaboration with the Computational and Medicinal Chemistry core of the Purdue Center for Cancer Research. To determine which domains are required for inhibition by these compounds, $USP7_{FL}$, $USP7_{CD}$, $USP7_{CD}$–H1–5, $USP7_{CD}$–H4–5, and TRAF–$USP7_{CD}$ were evaluated for inhibition. Interestingly, it was found that no inhibition could be measured without HUBL domains 1–3 present. With this finding, the mode of inhibition of these compounds was determined to be partial Non–Competitive, suggesting these inhibitors act via a novel allosteric inhibition mechanism. Many attempts were made to determine where these compounds bind to USP7 using both kinetic and biophysical approaches, however, no conclusive results were obtained. When combined, these observations support the importance of having an intact H1–3 for the intramolecular activation of USP7 and support that the novel allosteric inhibitors developed require H1–3 for inhibition. Recent x–ray structures of a USP7 truncation of $USP7_{CD}$–H1–3 may aid in identification of a more precise location for the design of inhibitors that target H1–3 domains.

CHAPTER 1. INTRODUCTION

1.1 Ubiquitination Cycle

1.1.1 Ubiquitination

Over the past 10 years our interest and understanding of the post–translational modification process of ubiquitination has greatly increased. The significance ubiquitination in maintaining cellular homeostasis by the regulation of many delicate cellular processes, including the cell–cycle, membrane protein trafficking, endocytosis, and apoptosis has become increasingly clear [1–4]. The ubiquitin protein itself is a conserved 76–residue polypeptide that folds into the highly stable β–grasp fold. The conjugation of ubiquitin to a protein substrate requires three enzymes: the ubiquitin–activating enzyme (E1 ligase), the ubiquitin–conjugating enzyme (E2 ligase), and the ubiquitin–protein ligase (E3 ligase). The ubiquitination process begins with an E1 ligase activating the C–terminus of a ubiquitin molecule through an ATP–dependent process forming a covalently bound intermediate between the active site of an E1 ligase ubiquitin (Figure 1.1). The activated ubiquitin is then transferred from an E1 ligase to the active site residue of the E2 ligase. The E3 ligase is responsible for recruiting the specific protein substrate for ubiquitination. Together the E2 and E3 ligases conjugate the ubiquitin to the protein by the formation of an isopeptide bond between the ϵ–amino of a specific lysine residue of the protein substrate and the C–terminus of ubiquitin [4, 5]. This process is reversed by deubiquitinating enzymes described in Section 1.2

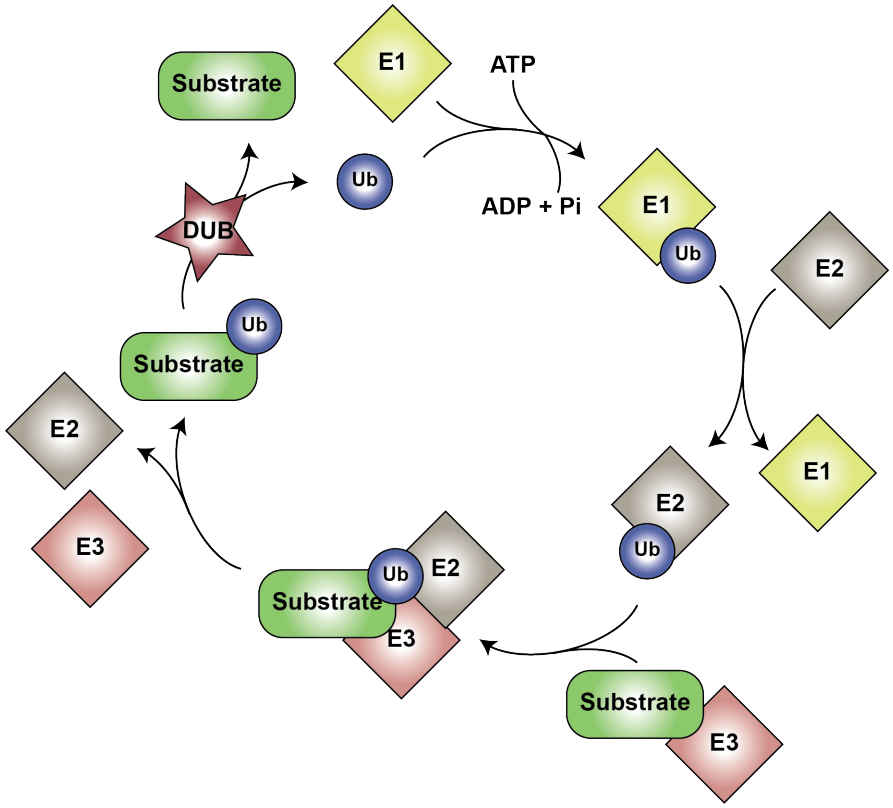


Fig. 1.1. **Ubiquitination Cycle.** Ubiquitination is a three step process. The C–terminus of a ubiquitin molecule is activated by the E1 Ligase through an ATP dependent process. The activated ubiquitin is then transferred to the E2 Ligase. The E3 Ligase is responsible for substrate recognition, and, together with the E2 Ligase, the activated ubiquitin molecule will be conjugated to the recruited substrate through the formation of an isopeptide bond between the C–terminus of ubiquitin and a specific lysine of the substrate. This process may be repeated numerous times to build ubiquitin chains from the lysines within ubiquitin itself.

1.1.2 E1, E2, and E3 Ligases

The cascade of E1, E2 and E3 ligases involved in the ubiquitination process serves to regulate and add specificity to the process. The human genome encodes for only 2 E1 ligases and 37 E2 ligases. In contrast, over 600 E3 ligases are encoded in the genome. With such a small number of E1 and E2 ligases, there can be minimal specificity towards specific protein substrates in the preparation of the ubiquitin molecule.

Instead, the specificity of the ubiquitination process is primarily the responsibility of the E3 ligases. The E3 ligases recognize both the E2 ligase and the protein substrate requiring ubquitination. There are two main types of E3 ligases, the HECT–type and the RING–type. The HECT–type ligases function alone as a single E3 ligase molecule in a two step process. The HECT–type E3 ligases form a covalent bond intermediate with the activated ubiquitin prior to transferring it to a protein it has recruited. Alternatively, the RING–type E3 ligases form multi–subunit complexes that can bind both the E2 ligase and a recruited protein substrate at once. With this ability, the RING–type E3 ligases transfer the activated ubiquitin molecule directly from the E2 ligase to the recruited protein substrate without forming a covalent intermediate. As a modular system, the RING–type E3 ligases can assemble to recognize different E2 ligases and different substrates in response to signals within the cell [6].

1.1.3 Ubiquitin Chains

A protein substrate can be modified by a single ubiquitin protein (mono-ubiquitination) or a chain of ubiquitin molecules (poly-ubiquitination) [3]. Ubiquitin chains are formed by the conjugation of the activated C–terminus of a ubiquitin molecule to the ϵ–amino of a lysine residue of a ubiquitin molecule already conjugated to the substrate [5]. Ubiquitin has seven different lysine residues from which poly–ubiquitin chains can be conjugated, Lys^{6}, Lys^{11}, Lys^{27}, Lys^{29}, Lys^{33}, Lys^{48}, Lys^{63}, in addition to the amino group of the N–terminal methionine, a modification referred to as linear ubiquitin [3]. The variety of ubiquitin chains that can be formed and the downstream consequences of the specific chain adds complexity to ubiquitin signaling that is lacking in other posttranslational modifications, such as phosphorylation. The outcomes of Lys^{48} and Lys^{63} poly–ubiquitination are the best understood thus far. Lys^{48} chains commonly target protein substrates for proteasomal degradation, while Lys^{63} chains are associated with enzyme activation [3]. Compared to Lys^{48} and Lys^{63}, an understanding of the physiological responses to the other ubiquitin chain

types is lacking. However, recent work has shed some light on the importance of Lys^{11} and Lys^{33}. The Lys^{11} linkage has been described as a degradation signal, similar to Lys^{48} [7]. Lys^{33} chains have been shown to be involved in enzyme activation, similar to Lys^{63} [7, 8].

1.2 Deubiquitinating Enzymes (DUBs)

Deubiquitinating enzymes (DUBs) counteract ubiquitination by cleaving the isopeptide bond between the ϵ–amino group of the protein substrate and the C–terminal group of ubiquitin. Approximately 98 human DUBs have been identified by sequence and structural similarity and organized based on these similarities into six families Figure 1.2: the Ubiquitin C–terminal Hydrolases (UCH), the Ubiquitin Specific Proteases (USP), the Machado–Joseph Disease Protein Domain Proteases (MCPIP), the Ovarian Tumor Proteases (OTU), the Josephine and Metalloenzymes (JAMM), and the monocyte chemotactic protein–induced protein family (MCPIP), as displayed in Figure 1.2 [1]. The members of these families are cysteine proteases, with the exception of the JAMM family, which is composed of metalloproteases. While they all perform the deubiquitinating function, structural studies have elucidated high variability within the catalytic domain between the families [1, 9].

DUBs aid in maintaining cellular homeostasis by six general processes. (1) Ubiquitin is translated as a single polypeptide chain, which requires DUBs to process the chain and release the individual ubiquitin molecules for conjugation to protein substrates by the ligase cascade. (2) When a protein is ubiquitinated to signal for degradation by the proteasome, DUBs will remove the ubiquitin tag from the substrate prior to degradation of the protein by the proteasome so as to recycle the ubiquitin molecules for future use. (3) DUBs can prevent protein degradation by the 26S proteasome by removing ubiquitin from targeted proteins. (4) The ubiquitination of a protein can act as an on/off switch for that protein's function. By removing ubiquitin from these types of proteins DUBs regulate the activity of the protein. (5)

DUBs can edit ubiquitin chains and hence signaling by altering the ubiquitination pattern on a particular protein. (6) A ubiquitin chain on a DUB binding partner can act as a signal to attract a DUB. This interaction can modulate the activity of the DUB or binding partner as a result [10].

DUBs recognize their substrates either by the substrate itself or by the specific ubiquitin chain that decorates the substrate [10, 11]. As previously discussed in Section 1.1.3, ubiquitin chains can be built out from seven different lysine residues within the ubiquitin molecule (Lys^{6}, Lys^{11}, Lys^{27}, Lys^{29}, Lys^{33}, Lys^{48}, Lys^{63}) in addition to the amino group of the N–terminal methionine. [3] The resulting ubiquitin chains extending from these different lysines assume different topologies that DUBs can recognize [10, 11]. Many DUBs are promiscuous and can process all ubiquitin chain linkages to some extent, in which case specificity is most likely inferred by the ubiquitinated substrate itself rather than the ubiquitin chain [10,11]. In some cases, however, DUBs can process only certain ubiquitin chains that they recognize directly. In both cases, DUBs can cleave the ubiquitin chains either by endo– or exo–trimming [10,12]. For endo–trimming, the DUB cleaves an isopeptide bond located within the ubiquitin chain, not the bond that links the ubiquitin molecule directly to the substrate or the last ubiquitin molecule in the chain. For exo–trimming the DUB cleaves the isopeptide bond between the substrate and the ubiquitin chain or the last ubiquitin molecule of the chain. Furthermore, some DUBs are capable of processing a mono–ubiquitinated substrate, while others are not and require a ubiquitin chain to perform catalysis [10, 12].

1.2.1 Ubiquitin Specific Proteases (USPs)

The USP family is the largest of the DUB families with around 60 members. The USP family is the focus of the work described within this dissertation. USPs are multi–domain enzymes that can range in size from 40 kDa, in the case of USP12, up to 400 kDa, in the case of USP34 [1]. The USP catalytic domain assumes a papain–

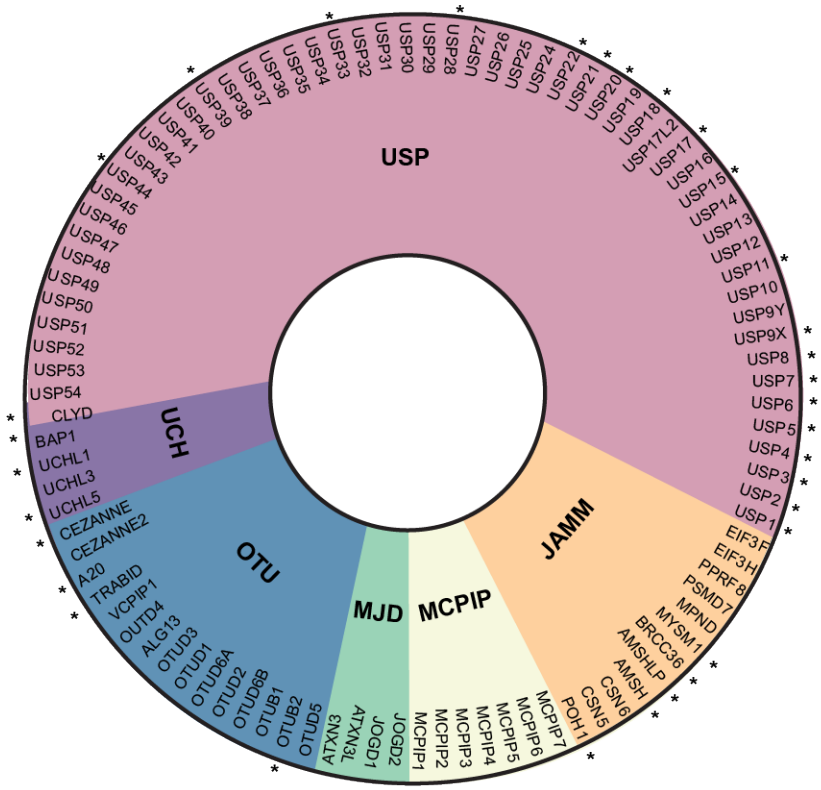


Fig. 1.2. **Classification of human DUBs** There are six families of human DUBs represented by the various colors: USPs (pink), UCH (purple), OTU (blue), MJD (green), MCPIP (yellow), and JAMM (orange). The known individual members of each family are listed within the appropriate section. The asterisks denote DUBs that have been characterized to be upregulated in cancer. Figure adapted from data summarized in [1, 13].

like fold (Figure 1.3A) that is structurally conserved throughout the family [1, 4]. The papain–like fold resembles a right hand with the fingers, thumb, and palm motifs (Figure 1.3A) [1]. Within the catalytic domain, there are two regions of high sequence similarity, which are referred to as the Cys (green) and His (cyan) boxes (Figure 1.3B). These conserved boxes are situated between the thumb and palm motifs and contain residues essential for catalysis including the catalytic triad that is composed of the catalytic cysteine (Cys box), and the supporting residues histidine and aspartate (His box) (shown as orange sticks in Figure 1.3B) [4, 9, 14].

Structural studies and comparisons between the USP family members have elucidated internal mechanisms the USPs utilize for regulation. One mechanism by which USPs are internally regulated is by the misalignment of their catalytic triad. The catalytic triad within the unligated structure both of USP1 and USP7 are positioned in a non–functional state [15–17]. However, ubiquitin–bound structures show the catalytic triad realigns into a functional state, suggesting that only when ubiquitin is bound is the USP functional [10, 16, 17]. All USPs contain loops within their catalytic domain referred to as blocking loops 1 and 2. However, the sequence and length of these loops are not conserved within the family. In the case of USP8 and USP14, these blocking loops physically occlude the active site and prevent catalysis (Figure 1.3C) [10,18,19]. The mechanism by which these loops move to allow ubiquitin binding is still being defined. Finally, while not clearly defined as a regulatory mechanism, a loop termed the "switching loop" has been described in USP4 and USP7 (Figure 1.3C) to be important for the allosteric activation of catalytic activity [20, 21].

What makes each USP unique are the domains that flank the conserved catalytic domain. These domains commonly include Ubiquitin–Like domains (UBL), Zinc–finger motifs, and Ubiquitin–interacting domains [9, 10]. These domains are often responsible for substrate recognition, protein–protein interactions, or even the regulation of USP activity [12,17,19,20]. Of particular interest for this dissertation are the UBL domains. UBL domains share the common β–grasp fold of ubiquitin but lack the C–terminal glycine required for conjugation. UBL domains are found throughout nature and are surprisingly commonplace in USPs [22, 23]. Thus far, at least 16 members of the USP family have been identified as containing UBL domains, and it appears that the role of these UBL domains differs from USP to USP. The first USP associated UBL domain was discovered within USP14. This UBL domain appears to be required for recruiting and associating USP14 with the proteasome [19]. USP4 also contains a UBL domain that was originally thought to have a role in inhibiting the catalytic activity of USP4 by blocking the active site, however this finding has yet

to be validated [21]. In the case of USP7, the UBL domains have a role in catalysis as well as serving as a binding platform for protein–protein interactions [20, 22].

1.2.2 Catalytic Mechanism of Ubiquitin Specific Proteases

The catalytic mechanism of USPs is illustrated in Figure 1.4. The USP catalytic triad is composed of a cysteine that acts as the nucleophile, a histidine that serves as a general acid–base, and an aspartate that stabilizes the histidine to promote the deprotonation of the cysteine residue. Within the unbound state of the USP, the histidine deprotonates the cysteine residue to generate the reactive thiolate (E). When the USP binds substrate, the ES complex is formed. The reactive thiolate will nucleophilically attack the isopeptide bond between the ubiquitin molecule and the protein substrate forming a tetrahedral intermediate (FP). The tetrahedral intermediate is unstable and quickly collapses by general acid assisted protonation by the histidine, releasing the product from the ubiquitin forming a thiol–ester intermediate (F). An active site water molecule then acts as a nucleophile to attack the carbonyl carbon of the thiol–ester intermediate forming a second tetrahedral intermediate (FQ). Again, the tetrahedral intermediate collapses through the general acid assisted protonation by the histidine, the ubiquitin molecule dissociates, and the enzyme returns to its initial state (E + Q).

1.3 USPs and Cancer

USPs regulate many sensitive pathways, as described in Section 1.2. As a result, the USPs are commonly associated with oncogenic pathways and are often a driving force of cancerous growth [13, 24]. The USPs and other DUBs upregulated in cancers are denoted in Figure 1.2 with an asterisk. The role each USP has in cancer is not always clear. Often times an individual USP can have both oncogenic and tumor suppressive traits, depending upon the specific pathway of involved or the tissue being studied [13]. By characterizing individual USPs and developing a more comprehensive

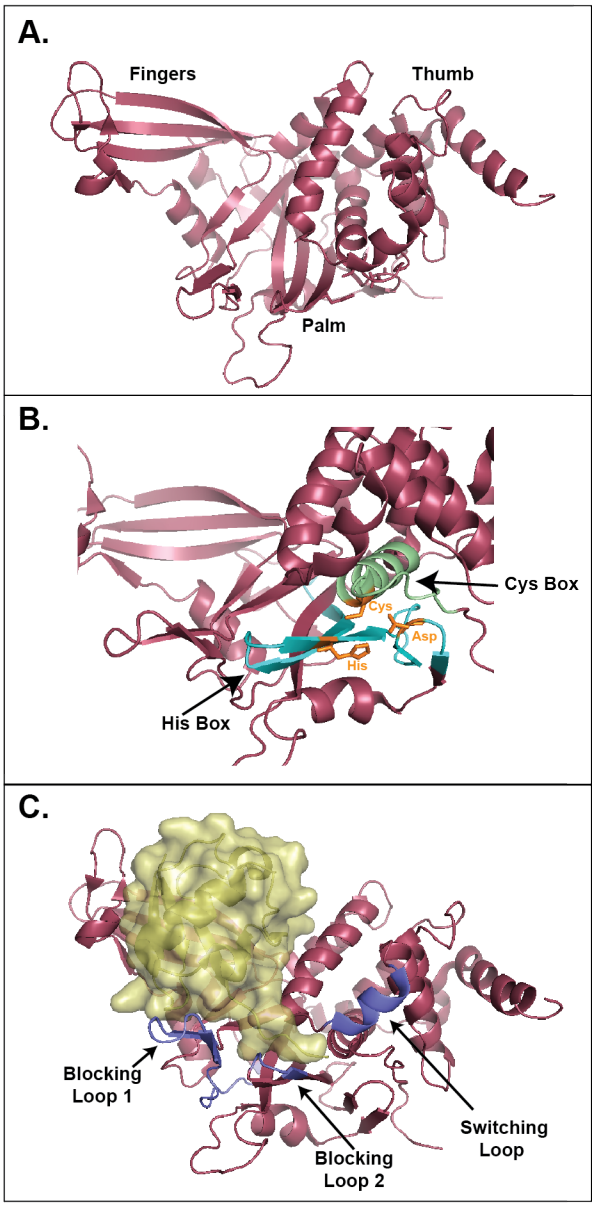


Fig. 1.3. **Conserved USP Catalytic Domain A.** Conserved papain–like fold present in USP catalytic domains. The fingers, thumb, and palm motifs are as labeled. Structure is of USP7 catalytic domain, PDB: 4M5W. **B.** Highlighted sequence–conserved Cys (green loop) and His (cyan loop) boxes. The catalytic triad is shown in orange: Cysteine, contained in the Cys box loop, and the Histidine and Aspartate, contained within the His box loop. **C.** Highlighted structural features blocking loops and switching loop in purple. The blocking loops are present in all USPs but vary in sequence and length. The switching loop has only been described as an important feature of USP4 and USP7.

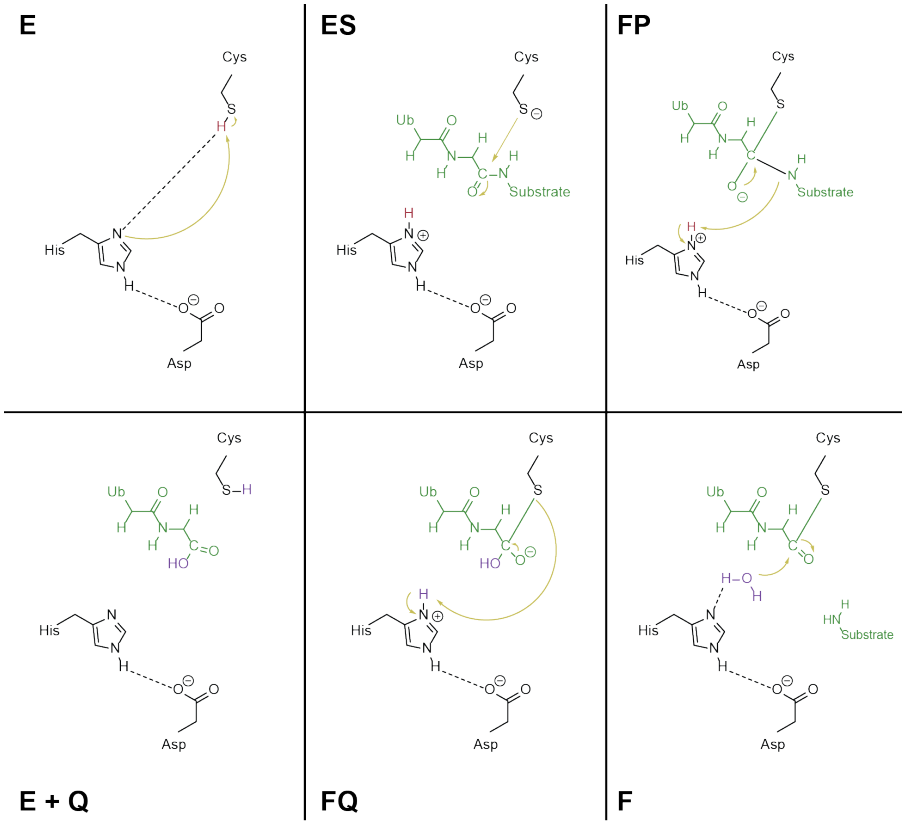


Fig. 1.4. **Catalytic Mechanism for Ubiquitin Specific Proteases** To assume the proper ionization state for catalysis (E), the active site histidine of the USPs catalytic triad acts as a general base to extract a proton from the catalytic cysteine. In the ionized state, the enzyme can form an ES complex with an ubiquitinated substrate. The thiolate ion then attacks the carbonyl of the last glycine residue of the ubiqutin molecule to form the tetrahedral intermediate FP. The unstable oxyanion drives the protonation of the leaving group product (the now deubiquitinated substrate) by the catalytic histidine to form the acyl enzyme (F). An activated water molecule then attacks the newly formed carbonyl to again form an oxyanion tetrahedrial intermediate and re–protonating the active site histidine (FQ). By general acid protonation of the thiolate by the histidine the tetrahedral intermediate collapses which regenerates free enzyme (E+Q).

understanding of how they regulate specific pathways, we can develop an understanding of how to target these enzymes for inhibition. The work described within this dissertation pertains to characterizing the enzymatic activity of two USPs associ-

ated with promoting cancer when misregulated, USP17 (Section 1.3.1) and USP7 (Section 1.3.2).

1.3.1 USP17

USP17 was originally identified as a member of the murine hematopoietic specific genes as DUB3 [25]. It is a 58 kDa enzyme with the catalytic domain located near the N–terminus and two predicted hyaluronan binding motifs at the C-terminus (Figure 1.5A) [25, 26]. No structural data for USP17 is currently available, however sequence alignments of USP17 with other well studied USPs (USP8, USP7, and USP14) reveal that USP17 contains its catalytic triad within the conserved Cys and His boxes as described in Section 1.2.1 (Figure 1.5B, catalytic residues highlighted in yellow). The activity of USP17 is closely associated with the cell–cycle. Cellular expression of USP17 is cytokine–inducible and is required for cell–cycle progression through the G_1–S and G_2–M checkpoints. If USP17 is expressed at different times during the cell–cycle besides these checkpoints, cellular proliferation ceases [25, 27]. Characterization of USP17 function has thus far been based primarily on cell–based studies and has led to the identification of three protein substrates of USP17: Ras Converting Enzyme 1 (RCE1), the Cell Division Cycle 25A (CDC25A) phosphotase, and a component of the Sin3 co–repressor complex SDS3 [28,29]. While under normal conditions the overexpression of USP17 ceases proliferation, work by McFarlane and coworkers described that the persistent overexpression of USP17 applies continuous stress on the cell–cycle resulting in cancerous proliferation in both breast and prostate cancers [25,27]. The effects of overexpression of USP17 and its effect on the substrates RCE1 and CDC25A have been elucidated and are discussed in section Section 1.3.1.

USP17 and RCE1

In the Ras pathway, USP17 has the function of regulating RCE1. RCE1 is an endoplasmic reticulum, integral transmembrane protease responsible for processing

C–terminal -CAAX motif of Ras, which is required for isoprenylation. The removal of the "AAX" residues exposes a cysteine, which undergoes further modification (methylation and palmitoylation) for localization of Ras to the plasma membrane where Ras is activated for signaling through the MAP kinase pathway. RCE1 is activated only when ubiquitinated by Lys^{63} chains. In a healthy cell, USP17 regulates this pathway by removing the ubiquitin from RCE1 rendering it inactive for proper Ras processing [28, 30]. It is unclear at this time how the upregulation of USP17 disrupts this pathway leading to cancer, but it has been hypothesized that due to the innappropriate expression of USP17, Ras does not translocate to the plasma membrane appropriately. A build up of Ras within an intermediate state leads to prolonged stress within the cell and further impairs proper cell–cycle signaling [27].

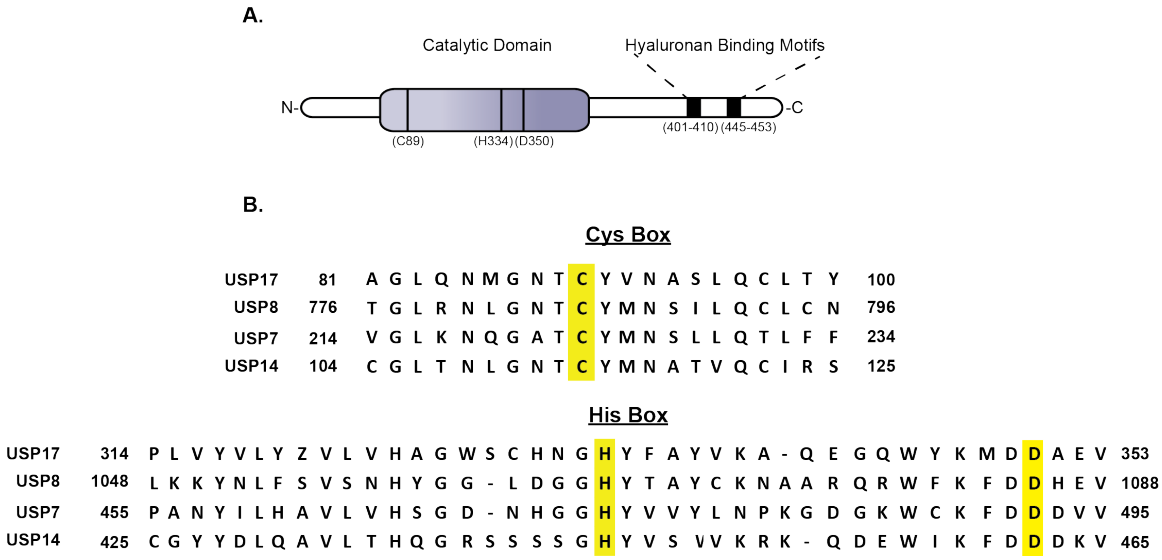


Fig. 1.5. **USP17 Domain Organization and Sequence Alignment A.** Schematic displaying the domains of USP17 identified thus far. **B.** A sequence alignment of the conserved Cys and His boxes of USP17 against USP8 (GenBank: 001122082), USP7 (GenBank: 003461), and USP14 (GenBank: 005142.1). Catalytic triad residues denoted by yellow highlighting.

USP17 and CDC25A

USP17 has been shown to have an important role in controlling the cell–cycle through the deubiquitination of CDC25A [29]. CDC25A is a cell–cycle phosphatase required for cell–cycle progression at the transition checkpoints and for response to DNA damage [29, 31]. CDC25A dephosphorylates the CDK2–Cyclin complex for transition through the cell–cycle checkpoints. USP17 is responsible for stabilizing CDC25A by removing Lys_{48} conjugated ubiquitin chains during the the transition checkpoints. Under normal conditions, CDC25A is ubiquitinated for degradation in a timely manner after transition through cell–cycle checkpoints (Figure 1.6A). The upregulation of USP17 leads to increased deubiquitination and inappropriate stabilization of CDC25A [29, 31]. Continuous upregulation of CDC25A stresses cell–cycle progression and leads to DNA damage and instability. Cells that experience persistent DNA damage often become cancerous (Figure 1.6B). The over–expression of USP17 has been attributed to promoting breast cancer and prostate cancer and is capable of inducing cellular transformation [25, 27, 31, 32]. The partial knock down of USP17 by siRNA was shown to reduce growth of breast tumor xenographs by 50% [29].

1.3.2 USP7

One of the most well studied DUBs is the ubiquitin specific protease 7 (USP7), also known as the herpes virus-associated ubiquitin-specific protease (HAUSP). X-ray crystal structures of each domain of USP7 are available as individual proteins, however, a full–length structure has yet to be determined. USP7 comprises of seven domains (Figure 1.7); a TNF receptor associated factor (TRAF) domain, a catalytic domain, and five ubiquitin-like (UBL) domains [19, 20, 33]. The N–terminal TRAF domain is proposed to be responsible for substrate binding and specificity [33]. The catalytic domain maintains the conserved papain–like fold seen throughout the USP family, (Figure 1.3) [19]. Interestingly, the purified catalytic domain is 120–fold less

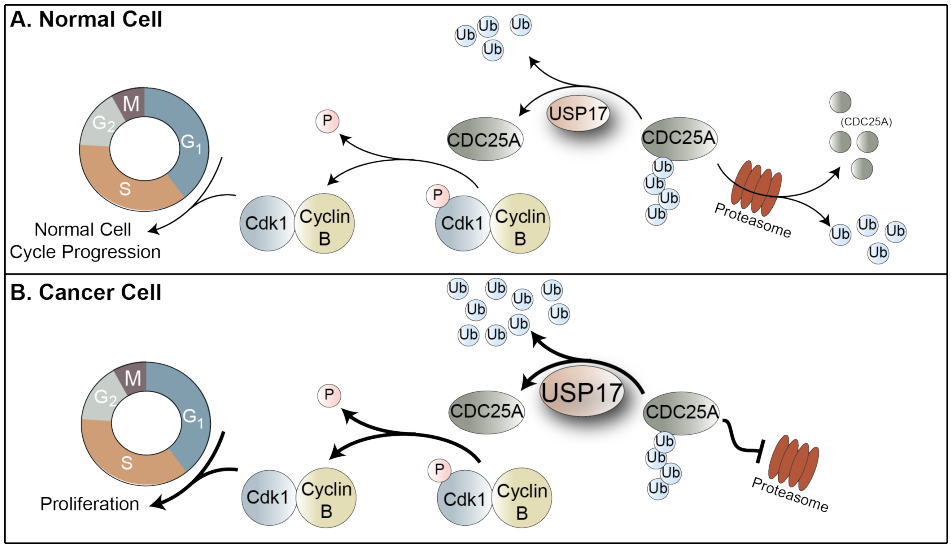


Fig. 1.6. **The role of USP17 in the CDC25A pathway A.** Under normal conditions, USP17 deubiquitinates CDC25A preventing the degradation of CDC25A by the proteasome. Stabilized CDC25A dephosphorylates the Cyclin–CDK2 complex for progression of the cell cycle through the G_1–S checkpoint. **B.** Under cancerous conditions (such as breast cancers), USP17 is upregulated which increases the amount of CDC25A present in the cell. This perturbation leads to increased progression of the cell–cycle and unregulated cellular proliferation.

active than that of the full–length enzyme as measured with the fluorogenic substrate ubiquitin–7-amino–4-methylcoumarin (Ub–AMC) [34]. The poor activity of the isolated catalytic domain has been attributed to the observation that the catalytic triad is misaligned in the unbound structure of USP7 (PDB: INB8 and 4M5W) presumably making it catalytically inefficient [15,16]. However, in a crystal structure of the catalytic domain bound to ubiquitin aldehyde (PDB: 1NBF), the catalytic triad is properly aligned for catalysis, suggesting that a rearrangement of the catalytic triad by binding of the substrate may be required for catalysis, as mentioned in Section 1.2.1 [16]. Furthermore, as discussed in Section 1.2.1, a comparison of the unbound structure of the catalytic domain to the ubiquitin aldehyde–bound structure shows that significant movement of blocking loops 1 (residues 408-429, BL1) and 2

(residues 459-462, BL2), as well as the switching loop (residues 285-291), are required to unblock the active site for the binding of ubiquitin–aldehyde [15, 19].

At the C–terminus of USP7 are five consecutive UBL domains referred to as the HAUSP ubiquitin-like (HUBL) domains [20]. The sequence identity of the UBL domains relative to each other and to ubiquitin is very low, and as a result, UBL domains are often only identified by determining their structures [22]. This remains true for the HUBL domains where the sequence identity among them is only 3 to 15%, or to ubiquitin, varying from 6 to 19% [20]. The significance of the HUBL domains in regulation of USP7 catalysis is slowly being defined and it is now clear that the HUBL domains, specifically HUBL domains 4 and 5, are required for full activity of USP7 [20, 34]. However, a detailed mechanism of how interactions with HUBL domains 4 and 5 influence USP7s activity has yet to be fully elucidated [20] and is a main topic of this dissertation.

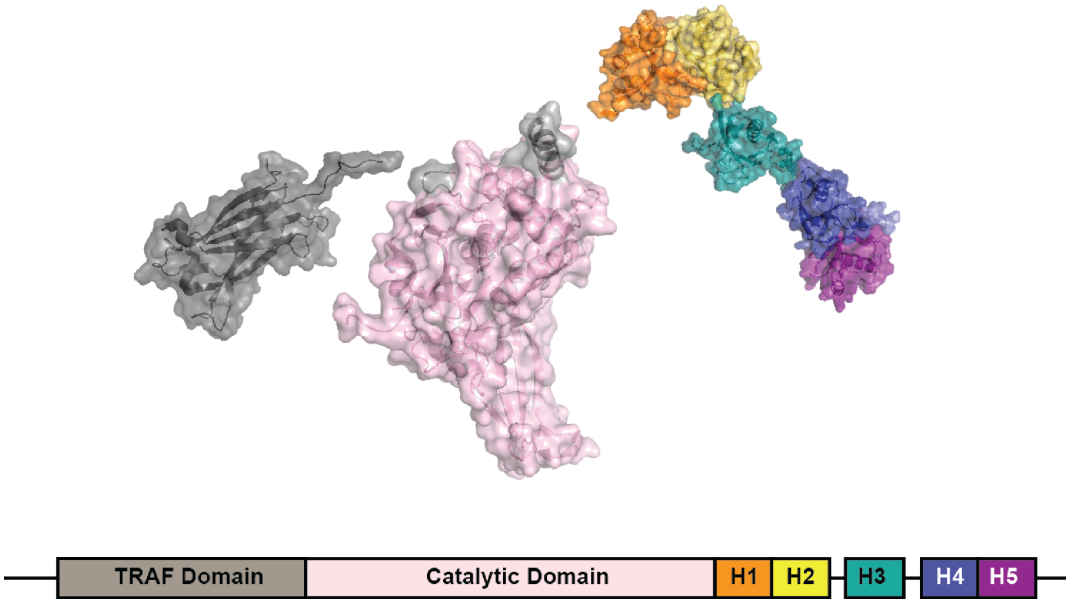


Fig. 1.7. **Domains of USP7** USP7 is comprised of seven domains: the N–terminal TRAF domain (grey), the conserved catalytic domain (pink), and five UBL domains (rainbow).

USP7 has many defined substrates involved in cell–cycle and growth signaling pathways. These substrates include tumor suppressors PTEN, FOXO4, and p53, as well as the E3 ligase ICP0, which is important for herpesvirus infection [35–39]. The upregulation of USP7 leads to many types of cancers by perturbing the balance of these signaling pathways. Cancers associated with USP7 upregulation include prostate, breast, glioma, lung, multiple myeloma, colon, and neuroblastoma [35, 36, 40–43].

USP7 and MDM2/p53

The most well studied substrates of USP7 are the tumor suppressor p53 and E3 ligase MDM2. Both substrates are associated with multiple myeloma. Recent advancements in the understanding of mulitple myeoma carcinogenesis have shown that the tumor suppressor p53 is silenced in the majority of cases. While p53 is mutated or deleted in about 50% of cancers, this is rarely the case in multiple myeloma [44]. p53 is directly regulated by ubiquitin–dependent degradation by MDM2. MDM2 autoubiquitinates itself for proteasomal degradation. In normal conditions, USP7 is responsible for deubiquitinating MDM2, thereby stabilizing MDM2 and allowing for continued proteasomal degradation of p53 [42]. Both USP7 and MDM2 levels are upregulated in multiple myeloma cells, which is effectively silencing p53 and resulting in unchecked oncogenic proliferation (Figure 1.8) [44, 45].

USP7 and PTEN/FOXO4

Mono–ubiquitination of the USP7 substrates phosphatase/tumor suppressor PTEN and tumor suppressor FOXO4, signals for their translocation to the nucleus [35, 36]. Within the nucleus both PTEN and FOXO4 function as transcriptional regulators for tumor suppression [35, 36]. As depicted in Figure 1.9, USP7 is responsible for the deubiquitination of both substrates, which leads to their export from the nucleus. The upregulation of USP7 in cancers, such as prostate and breast cancers, leads to

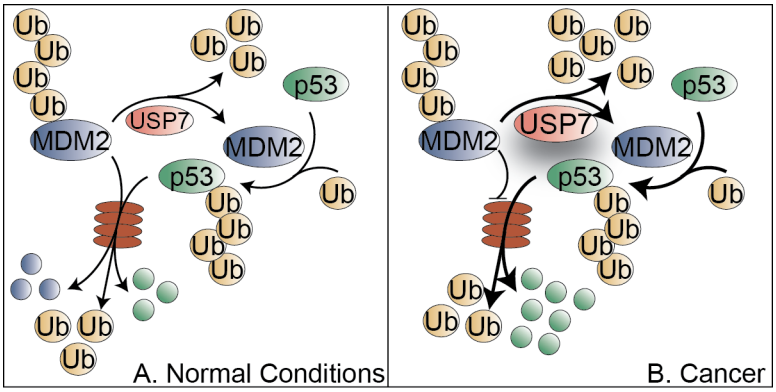


Fig. 1.8. **The role of USP7 on the p53/MDM2 pathway. A.** Under normal conditions, E3 ligase MDM2 ubiquitinates both p53 and itself for degradation by the proteasome. USP7 regulates the balance between MDM2 and p53 through deubiquitination. **B.** Under cancerous conditions, such as multiple myeoma, USP7 levels are upregulated, which leads to increased levels of MDM2. The effects of stabilizing MDM2 to such an extreme, essentially leads to the silencing of p53 and unchecked cellular proliferation.

the perturbation of these signaling pathways and increased export of these critical tumor suppressors from the nucleus. Without successful PTEN or FOXO4 signaling within the nucleus, inappropriate cellular proliferation results [35, 36].

1.4 Statement of Intent

The USP family of enzymes play an important role in maintaining homeostasis due to their regulation of a vast number of cellular processes including cell–cycle regulation, DNA damage repair, chromatin remodeling, or signaling pathways by deubiquitination [1]. Each USP family member shares a structurally conserved catalytic domain and is accessorized with other functional domains flanking the catalytic domain (Figure 1.3). It is the composition and function of these additional domains that makes the function of each USP family member unique. Almost one half of the USP family members identified thus far are upregulated in cancer (Figure 1.2). Of particular interest for the research described herein are USP17 and USP7 and the

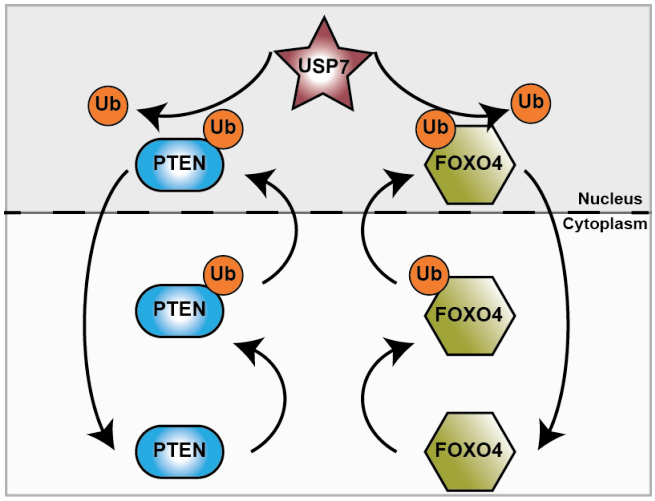


Fig. 1.9. **USP7 deubiquitination of PTEN and FOXO4.** Tumor suppressors PTEN and FOXO4 are both mono–ubiquitinated within the cytoplasm which signals for their translocation to the nucleus. USP7 maintains PTEN and FOXO4 levels by deubiquitination which subsequently signaling for their export from the nucleus. Increased levels of USP7 leads to the inappropriate export of these critical tumor suppressors from the nucleus and results in unregulated cellular proliferation and cancer.

role they play in promoting cancer. The overarching goal of this dissertation project is to characterize how each USP processes ubiquitin substrates as well as define the role of the additional domains within the enzyme. Understanding how each of these enzymes differs from the other members of the USP family may aid in the identification and development of specific small molecule inhibitors towards both enzymes. These inhibitors could further aid in the study of the enzymatic function of each USP and serve as a foundation for the development of anti–cancer therapeutics.

CHAPTER 2. ENZYMATIC CHARACTERIZATION OF RECOMBINANTLY EXPRESSED USP17

Parts of the data and text in this chapter have been included in the manuscript in preparation entitled "Steady–state kinetic studies reveal that the anti–cancer target Ubiquitin–Specific Protease 17 (USP17) is a highly efficient deubiquitinating enzyme"

2.1 Abstract

USP17 is a deubiquitinating enzyme that is responsible for the regulation of cell–cycle progression by stabilizing the Cell Division Cycle 25 A (CDC25A) phosphatase by deubiquitination. Upregulation of USP17 in numerous cancer phenotypes stabilizes CDC25A which leads to cellular proliferation and cancer, suggesting that an inhibitor of USP17 may have therapeutic value in the treatment of cancer. Towards this goal, we developed a robust expression, purification, and assay system for USP17 enabling its enzymatic and structural characterization and the identification of inhibitors of this potential drug target. USP17 was expressed in *E. coli* as inclusion bodies and then solubilized, refolded, and purified using affinity–chromatography followed by size–exclusion chromatography. Milligram quantities of highly pure, recombinant USP17 can be produced from a liter of culture, and the resulting enzyme is catalytically more efficient ($k_{\mathrm{cat}}/\mathrm{K_m}\bar{1}.5\ M^{-1}sec^{-1}$) than other human USPs studied to date. Analytical–size exclusion chromatography, analytical ultracentrifugation, and dynamic light scattering studies suggest that the quaternary structure of USP17 is a monomer. Steady–state kinetic studies show that USP17 efficiently hydrolyzes both ubiquitin-AMC ($k_{\mathrm{cat}} = 1.5\ sec^{-1}$ and $\mathrm{K_m} = 1.0\ \mu$M) and ubiquitin-rhodamine110 ($k_{\mathrm{cat}} = 1.8\ sec^{-1}$ and $\mathrm{K_m} = 2.0\ \mu$M) substrates. Ubiquitin chain cleavage assays reveal that USP17 efficiently cleaves di–ubiquitin chains with $Lys^{11}, Lys^{33}, Lys^{48}$ and

Lys^{63} linkages and tetra–ubiquitin chains with Lys^{11}, Lys^{48} and Lys^{63} linkages but is inefficient in cleaving di–ubiquitin chains with Lys^{6}, Lys^{27}, or Lys^{29} linkages or linear ubiquitin chains. The substrate specificity of USP17 is most similar to that of USP1, where both USPs display higher specificity than other characterized members of the USP family.

2.2 Introduction

Deubiquitinating enzymes (DUBs) reverse the process of ubiquitination by hydrolyzing ubiquitin from the protein substrate to that it is conjugated [1, 2]. Thus far, approximately 100 human DUBs have been identified in the human genome, and they are involved in regulating a number of cellular processes and disease states [1,2]. DUBs are classified into six families, of them, the largest family is the Ubiquitin Specific Protease (USP) family with roughly 60 members [1,2]. The USPs are multi–domain enzymes that can range in size from ∼40 kDa, in the case of USP12, up to ∼400 kDa, in the case of USP34 [1]. Each USP contains a catalytic domain which shares the highly conserved papain–like fold composed of a catalytic triad of cysteine, histidine, and aspartate residues [1, 2]. Other domains within the USPs are often important for protein–protein interactions and/or substrate recognition. A detailed review of the domains and their functions has been described elsewhere [2, 20].

USP17 was originally identified as a member of the murine hematopoietic specific genes as DUB3 [25]. Expression of USP17 within the cell is cytokine–inducible, and is required for cell–cycle progression through the $G_1 - -S$ and $G_2 - -M$ checkpoints. A well–characterized substrate of USP17 is the cell division cycle 25A (CDC25A) phosphatase. In a normal cell, CDC25A is responsible for dephosphorylation of Cdk1 and activation of the Cdk1–CyclinB complex for the progression of the cell–cycle (Figure 1.6 A) [29]. Cellular levels of CDC25A are regulated by the ubiquitin–proteasome system through Lys^{48} ubiquitin chains. During the $G_1 - -S$ and $G_2 - -M$ checkpoints, USP17 is expressed to deubiquitinate and stabilize CDC25A in order to

promote cell–cycle progression through these checkpoints [29]. In numerous cancer phenotypes, expression of USP17 is upregulated outside of these checkpoints which causes an excess of stabilized CDC25A that applies stress to the cell–cycle resulting in unregulated cancer cell proliferation (Figure 1.6 B) [27, 29, 46]. Other identified substrates of USP17 include the Ras converting enzyme 1 (RCE1) and the histone deacetylase dependent Sin3A co–repressor complex, SDS3, both of them also play a role in cell–cycle regulation [28, 30, 47].

USP17 is a 58 kDa protein that has its conserved catalytic domain near the N–terminus as well as two hyaluronan binding motifs predicted to reside within the C–terminal region (Figure 1.5A) [25, 26]. A sequence alignment of the catalytic regions of USP17 against three well studied USPs (USP8, USP7, and USP14) is shown in (Figure 1.5B). As with other USPs, there are three well conserved residues (Cys89, His334 and Asp350) that form the catalytic triad. Experimental binding studies by Ramakrishna and coworkers have shown the two predicted hyaluronan binding motifs bind to hyaluronan, a polysaccharide that is responsible for numerous cellular processes, including the regulation of cell division [26]. Many USPs are predicted to contain hyaluronan binding motifs, including the well–studied USP7, however the interactions between the USPs and hyaluronon itself is poorly understood [48]. In the case of USP17, the hyaluronan binding motifs have been shown to be important for the interaction of USP17 with SDS3 [49].

McFarlane and coworkers have shown that the persistent overexpression of USP17 applies stress to the cell–cycle which results in proliferation of both breast and prostate cancers [25, 27]. Combining this finding with the characterization of the role of USP17 in CDC25A–driven cellular proliferation by Pereg and coworkers, together these studies strongly suggest that inhibitors of USP17 activity may serve as anti–cancer drugs. While previous work has been able to define the role of USP17 *in celluo*, without the ability to recombinatly express and purify USP17, further characterization of USP17 as well as the identification of USP17 inhibitors is severely limited. This chapter describes a reliable and robust method for the expression and

purification of recombinant USP17 which enables the production of highly pure and active USP17 that is amenable to enzymatic and structural characterization as well as inhibitor identification.

2.3 Experimental Procedures

2.3.1 Design of Expression Plasmid for USP17 from *E. coli*

The human gene of *usp17* (GenBank: Q0WX57.2) was codon optimized for *E. coli* expression, synthesized, and inserted into the pET11a expression vector by BioBasic Inc. The coding region for USP17 was amplified using forward and reverse primers containing complementary sequences to the expression vector, pEV–L8, a modified pET30 vector. The primers used are listed 5' to 3' where the uppercase bases are complimentary to the pEV–L8 vector and the lower case bases are complimentary to the *USP17* gene: forward primer: GAGAACCTGTACTTCCAATCCAATatggaagacgatagcctgtac, reverse primer: cgtgacgaccagacagttATTTAACCTTCACCTATTGCCTAGGCTTAAGCT. The master mix for a single 25 μL PCR reaction is summarized in Table 2.1.

Table 2.1
PCR composition for the sub–cloning of the *usp17* gene

Component	Final Concentration	Volume
5x Phusion HF Buffer	1x	5 μL
10 mM dNTPs	0.2 mM	0.5 μL
50 μM Forward Primer	0.25 μM	0.125 μL
50 μM Reverse Primer	0.25 μM	0.125 μL
USP17–pET11a plasmid	50–100 ng	–
Phusion High–Fidelity DNA Polymerase	0.5 units	0.25 μL
Sigma–Water	N/A	to 25 μL

The PCR reaction was performed with a Bio–Rad MJ Mini Personal Thermal Cycler with the following reaction scheme:

- Initial Denaturation: 98 °C – 3 minutes
- Denaturation: 98 °C – 30 seconds
- Annealing Gradient: 58–64 °C – 30 seconds
- Extension: 72 °C – 60 seconds
- Repeat Denaturation – Extention Steps for 30 cycles
- Completion: 72 °C – 10 minutes
- Hold: 4 °C

The expression vector pEV–L8 was digested with restriction enzyme SspI in a 50 μL reaction (1x Buffer G (Promega), 1 unit of SspI (Promega), and 5 μg pEV–L8) incubated in a 37 °C water bath for one hour. The PCR amplicon was then combined in a 1:1 ratio (by concentration) with the digested the pEV–L8 vector by ligation–independent recombinant cloning using the linearized pEV–L8 vector, and XL1–Blue supercompetent cells. The resulting USP17 pEV–L8 expression plasmid sequence was confirmed at the Purdue University Genomics Core Facility, and was then transformed into *E. coli* BL21 (DE3) cells for protein expression.

2.3.2 Site–Directed Mutagenesis for USP17 Catalytic Mutants

The catalytic cysteine of USP17 was mutated to a serine within the pEV–L8 expression vector using the QuikChange site–directed mutagenesis protocol from (Stratagene). The primers are as follows 5' to 3' where the bold text base correspond to the mutation: forward: GGGCAATACCT**C**CTATGTTAACGCATCTCTGC, rev erse:CATAGGAGGTATTGCCCATGTTCTGCAGACC.

The catalytic cysteine to alanine mutant pEV–L8 vector was also made as described for the cysteine to serine mutant, however this construct was not utilized within the work described in this text. The mutagenesis primers are as follows 5' to 3' where the bold text bases corresponds to the mutation: forward: GGGCAATAC CGCCTATGTTAACGCATCTCTGCAATG, reverse: GTTAACATAGGCGGTA TTGCCCATGTTCTGCAGACC. The master mix for a single 25 μL PCR reaction is summarized in Table 2.2.

Table 2.2
PCR composition for site–directed mutagenesis of the *usp17* gene

Component	Final Concentration	Volume
5x Phusion HF Buffer	1x	5 μL
10 mM dNTPs	0.2 mM	0.5 μL
10 μM Forward Primer	0.5 μM	1.25 μL
10 μM Reverse Primer	0.5 μM	1.25 μL
USP17–pEV–L8 plasmid	50–100 ng	–
Phusion High–Fidelity DNA Polymerase	0.5 units	0.25 μL
Sigma–Water	N/A	to 25 μL

The PCR reaction was performed with a Bio–Rad MJ Mini Personal Thermal Cycler with an initial denaturation for 30 seconds at 98 °C. Sixteen cycles were completed of denaturation at 98 °C for 10 seconds followed by an annealing gradient of 60–73 °C for 30 seconds and an extension at 72 °C for three minutes. After the 16 cycles were complete, the reactions were held at 72 °C for 60 minutes. The PCR reactions were visualized on a 1% agarose gel, and, as all reactions produced product, the PCR reactions were pooled, and underwent a DpnI digest to remove the original template DNA. The DpnI digest mixture was prepared with 5 μL of 10x Tango Buffer (Thermo Scientific), 1 μL of DpnI Restriction Enzyme (Thermo Scientific), and 44 μL of PCR Product.

The reaction was run for 2.5 hours in a 37 °C incubator. The digested sample was transformed by electroporation into *E. coli* XL1–Blue cells, and streaked onto Kanamycin plates for growth overnight at 37 °C. A single colony was used from each mutation to innoculate a 2 mL LB culture supplemented with 50 μg/mL Kanamycin. The cultures were grown at 37 °C overnight with shaking at 200 rpm in an Infors Multitron shaker. The cells were harvested, and mini–prepped the following morning using the Qiagen Mini–Prep Kit (Cat# 27104) to isolate the mutant plasmid. The resulting plasmids were sequenced at the Purdue University Genomics Core Facility.

2.3.3 Design of Expression Plasmids for USP17 Truncations from *E. coli*

As depicted in Figure 2.1B. three truncations of USP17 were made to attempt to improve solubility and stability. These truncations were made by ligation–independent cloning as described for the subcloning of *usp17* from pET11a into pEV–L8 described above. The truncations and corresponding primers are displayed below. Only the USP17 complimentary region for each primer is displayed from 5' to 3' within Table 2.3

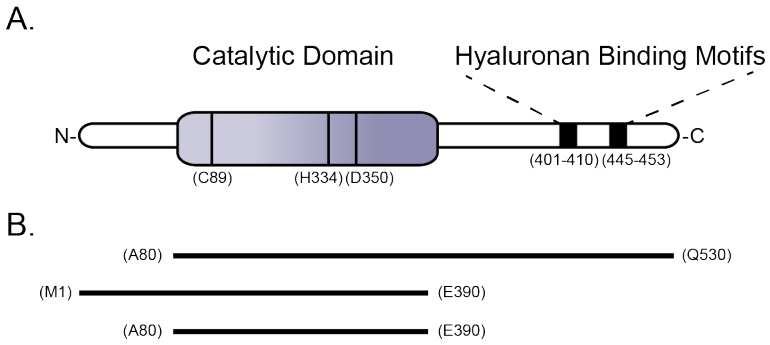


Fig. 2.1. **USP17 Truncations A.** Schematic displaying the domains of USP17 identified thus far. **B.** Outlines the three truncations of USP17 that were attempted to be expressed and purified.

The following sequence was added to the 5' side of the forward primers, and is complimentary to the pEV–L8 vector for homologus recombination: GAGAACCT-

Table 2.3
Primers for Truncations of USP17

Truncation Residues	Forward Primer	Reverse Primer
A80–Q530	gcaggtctgcagaacatg	ttgacagaccagcagtgc
M1–E390	atggaagacgatagcctgtac	ttcacgaccacgcga
A80–E390	gcaggtctgcagaacatg	ttcacgaccacgcga

GTACTTCCAATCCAAT. The following sequence was added to the 3' side of the reverse primers, and is complimentary to the pEV–L8 vector for homologus recombination: ATTTAACCTTCACCTATTGCCTAGGCTTAAGCT. The PCR mix for a single 50 μL reaction was composed as described in Table 2.4.

Table 2.4
PCR composition for truncations of USP17

Component	Final Concentration	Volume
5x Phusion HF Buffer	1x	10 μL
10 mM dNTPs	0.2 mM	1 μL
10 μM Forward Primer	0.25 μM	1.25 μL
10 μM Reverse Primer	0.25 μM	1.25 μL
USP17–pEV–L8 plasmid	50–100 ng	–
Phusion High–Fidelity DNA Polymerase	0.5 units	0.25 μL
Sigma–Water	N/A	to 50 μL

The PCR reaction was performed with a Bio–Rad MJ Mini Personal Thermal Cycler with an initial denaturation at 95 °C for three minutes. This initial step was followed by 30 cycles of the following: denaturation at 95 °C for 10 seconds, annealing at 49 °C for 30 seconds, and extension at 72 °C for 1.5 minutes. The reaction was

held at 72 °C to allow for completion for 10 minutes then held at 4 °C until further processing.

Each truncation PCR amplicon was then inserted into the pEV–L8 vector by ligation–independent recombinant cloning using linearized pEV–L8 vector, digested by restriction enzyme SspI, and XL1–Blue supercompetent cells as described above. The resulting USP17 pEV–L8 truncation expression plasmids were sequenced at the Purdue University Genomics Core Facility to confirm correct sequence, and was then transformed into *E. coli* BL21 (DE3) cells for protein expression.

2.3.4 Design and Amplification of Expression Plasmid for USP17 for Baculovirus Expression

To express within the baculovirus expression system, *usp17* (GenBank: Q0WX57.2) was codon optimized for *Spodoptera frugiperda 9* (*Sf9*) expression and synthesized by BioBasic Inc. The gene was inserted into the pVL–1392 vector modified with a 10His tag and a Green Fluorescent Protein (GFP) N-terminal tag by BioBasic Inc. The vector was transformed into *E. coli* XL1–Blue cells, and the cells were streaked onto LB agar containing 50 μg/mL Carbenicillin, and allowed to grow overnight at 37 °C. A single colony was used to innoculated 2 mL of LB media supplemented with 50 μg/mL Carbenicillin, then placed in an incubated Infors Multitron shaker at 37 °C shaking at 200 rpm and allowed to out grow for eight hours. Of the 2 mL culture, 25 μL of culture was then added to 25 mL of LB supplemented with 50 μg/mL of Carbenicillin in a 250 mL flask. The flask was placed in the incubated shaker at 37 °C with shaking at 200 rpm overnight. The cells were centrifuged with a Sorvall RC 6+ ThermoScientific Centrifuge in a Fiberlite F13–14 x 50cy fixed angle rotor at 3000 x g at 4 °C for 15 minutes. The expression plasmid was isolated from the cells by gravity flow using the Qiagen Midi–Prep Kit (Cat#: 12143). The isopropanol precipitation described in the protocol by Qiagen was performed in a chemical fume hood for sterility.

2.3.5 Expression, Solubilization, Refolding, and Purification of USP17 and USP17C89S from *E. coli*

Five liters of *E. coli* BL21 (DE3) cells harboring the USP17 pEV–L8 expression vector were grown from five separate one liter cultures of LB medium supplemented with 50 μg/mL Carbenicillin in 2 L Fernbach flasks. Shaking was performed at 37 °C and 200 rpm in an Infors Multitron shaker until an $A_{600} \sim 0.6$ was reached. Cultures were the cooled for 30 minutes at 4 °C, and USP17 expression was induced with 1 mM IPTG. Cultures were incubated at 25 °C with shaking at 200 rpms for an additional 18 hours.

E. coli cells were harvested by centrifugation with a Sorvall RC 6+ ThermoScientific Centrifuge in a Fiberlite F9–4x1000y rotor at 3000 x g at 4 °C for 20 minutes. The *E. coli* cells were resuspended in lysis buffer (50 mM Tris pH 7.5, 10 mM β ME, and 150 mM NaCl) supplemented with 50 μg/mL DNase and 200 μg/mL Lysozyme by adding 5 mL of lysis buffer per gram of cells. To lyse, the resuspended cells were sonicated on ice with a Branson Digital Sonifer at 60% with cycles of five second pulses and nine seconds of rest for a total of 12 minutes. The soluble lysate was separated from the inclusion body fraction by centriguation in the Sorvall RC 6+ ThermoScientific Centrifuge with the Fiberlite F13–14 x 50cy fixed angle rotor at 28,880 x g at 4 °C for 30 minutes. The resulting supernatant was decanted from the inclusion body pellet. The resulting inclusion body was washed 3 times by homogenization with 50 mL of lysis buffer containing 1% $Triton^{TM}X-100$. Between each wash, the sample was centrifuged at 28,800 x g for 10 minutes. A final wash was performed identically with unsupplemented lysis buffer.

Refolding of USP17

The washed inclusion body was then re–solubilized in 250 mL of re–solubilization buffer (6 M GuHCl, 50 mM CHES pH 9.5, and 10 mM βME) overnight in a sealed Nalgene bottle with a magnetic stir bar at room temperature on a stir plate with

rapid stirring at a rate of 600 rpm. The 250 mL of solublized inclusion body was refolded dropwise at room temperature into 5 L of refolding buffer (50 mM CHES pH 9.5, 10 mM βME, 0.7 M GuHCl, and 5% Glycerol) in a Nalgene bottle stirring on a magnetic stir plate at 600 rpms. Once all the protein had been refolded, the bottle was sealed, and the sample was incubated for 30 minutes stirring at 600 rpms. The refolded protein was then concentrated to 1 mg/mL in a volume of $\leq$ 500 mL using the Millipore Prep/Scale-Tangential Flow Filter concentrating system with a 1.5 L, 10,000 MWCO cartridge. The concentrated protein was centrifuged at 28,880 x g at 4 °C for 30 minutes, and filtered through a 0.45 μm filter to remove precipitant.

Nickel Affinity Chromatography

A 5 mL HisTrap Ni–NTA (GE Healthcare) chromatography column was equilibrated with Buffer A (50 mM CHES pH 9.5, 10 mM βME, 0.7 M GuHCl, 20 mM Imidizole, and 5% Glycerol). The filtered–refolded protein was then loaded onto the equilibrated HisTrap Ni–NTA column overnight at 4 °C with a flow rate of 0.5 mL/min.Unbound proteins were washed from the column using 15 column volumes of Buffer A that was applied at a flow rate of 2 mL/min.Bound proteins were eluted from the column using a 20 column volume gradient of 0-100% Buffer B (50 mM CHES pH 9.5, 10 mM βME, 0.7 M GuHCl, 450 mM Imidizole, and 5% Glycerol) that was applied at a flow rate of 2 mL/min. Fractions of five milliliters were collected throughout the gradient. The concentration of the protein in fractions was measured using the Microassay Bio-Rad Bradford Protein Assay. Peak fractions 6–14 containing USP17, as determined by SDS–PAGE (Figure 2.2A.) and specific activity, were pooled.

Gel–Filtration Chromatography

The pooled Nickel Affinity sample was concentrated to 7 mg/mL in a volume of $\leq$ 10 mL using ultrafiltration with a 30,000 Da MWCO membrane (Millipore).

The concentrated sample was injected onto a 300 mL Superdex 75 (GE Healthcare) SEC column equilibrated with Size Exclusion Buffer (50 mM CHES pH 9.5, 10 mM βME, 0.7 M GuHCL, and 5% Glycerol). The column was run at 1 mL/min and 5 mL fractions were collected continuously throughout. Peak fractions 15 and 16 containing USP17, as confirmed by SDS–PAGE (Figure 2.2B) and specific activity, were pooled, flash–frozen in liquid nitrogen, and stored at -80 °C until further use. Approximately 4.2 mg of pure protein can be obtained from one liter cell culture. A summary of the percent enzyme yield, total activity, and fold–purification after each chromatography step is summarized in Table 2.5. The catalytic mutant, USP17C89S, was purified from 1 L of culture identically (Figure 2.3) except the activity assays were omitted since the enzyme was inactive.

Table 2.5
Purification summary of USP17 from 5 L culture of *E. coli* BL21-DE3

Sample	Protein (mg)	Total Activity Units	Specific Activity (Units/mg)	Fold Purification	% Yield
Soluble Lysate	870	5.37	0.006	N/A	N/A
Refolded Protein	494	14.01	0.028	1	100
Nickel Pool	72	3.84	0.123	4	13
Superdex 75	21	3.12	0.147	5	10.5

2.3.6 Expression and Refolding Tests for Truncactions of USP17 from *E. coli*

For each construct, a 2 mL LB culture supplemented with 50 μg/mL Kanamycin was grown overnight at 37 °C with shaking at 200 rpm in an Infors Multitron shaker. The next morning, 125 μL of each culture was used to innoculate a 5 mL LB culture supplemented with 50 μg/mL Kanamycin. Each culture was incubated at 37 °C with

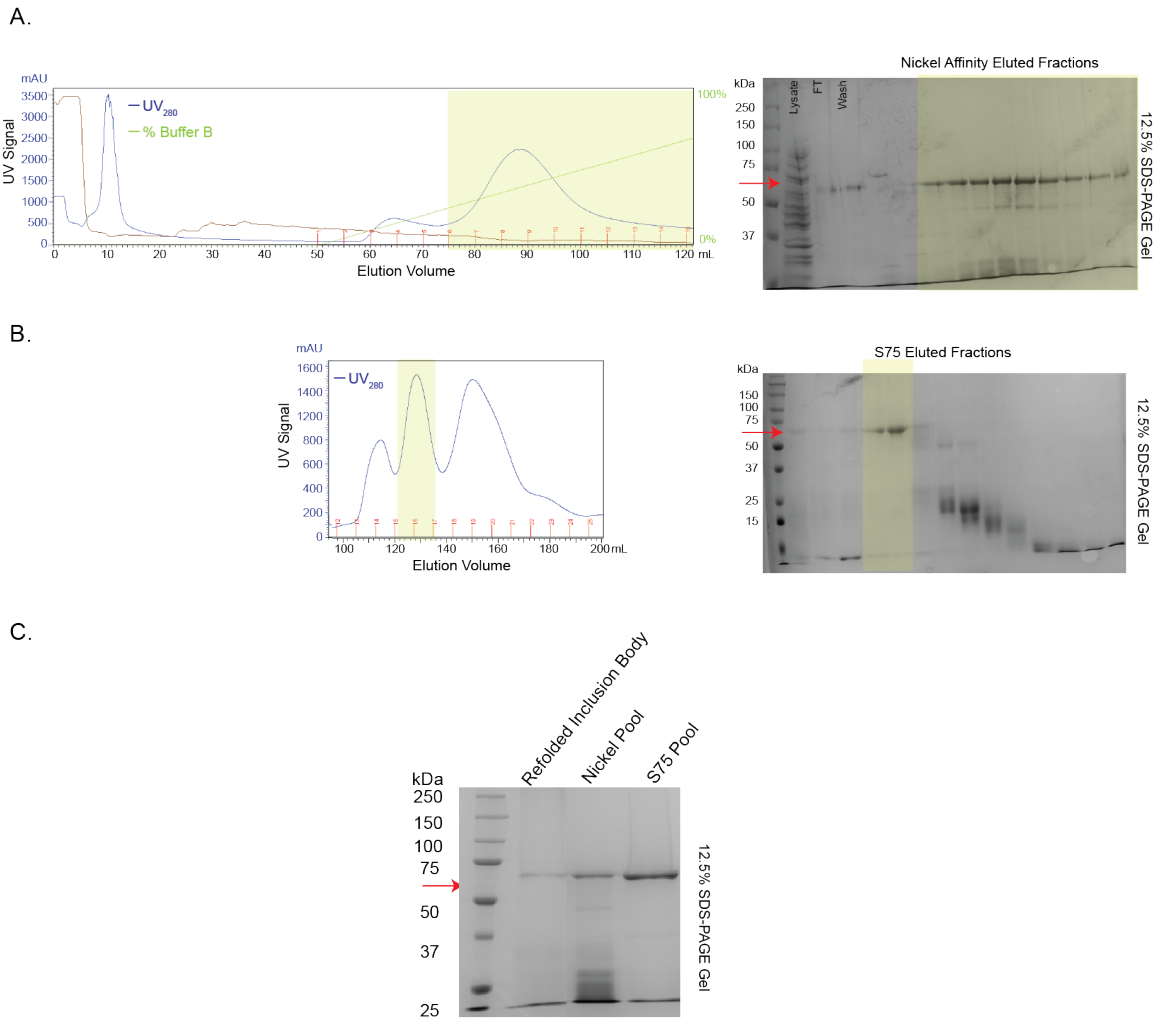


Fig. 2.2. **Purification of USP17. A.** *Left*–Elution profile from Hi–Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*–12.5% SDS–PAGE analysis of the eluted fractions, of which fractions 6–14 were pooled. Red arrow indicates expected size of USP17. Molecular weight marker sizes as indicated. **B.** *Left*–Elution profile from Superdex 75 column. *Right*– 12.5% SDS–PAGE analysis of eluted fractions, of which fractions 15 and 16 were pooled. *C.* A summary 12.5% SDS–PAGE analysis of each step of the purification of USP17.

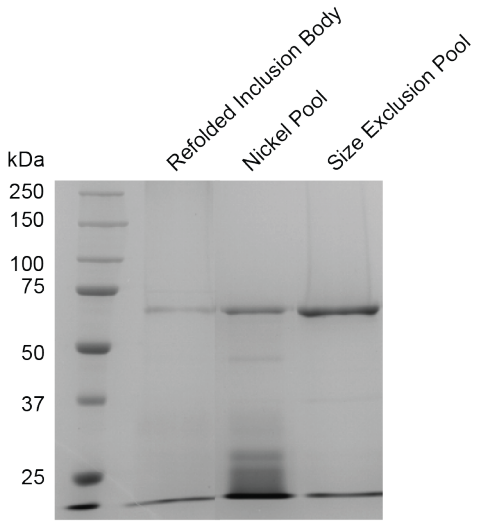


Fig. 2.3. **Purification of USP17C89S. A.** A summary 12.5% SDS–PAGE analysis of each step of the purification of USP17C89S.

200 rpm shaking until the A_{600} reached 0.6 at that time 1 mM IPTG was added to the culture to induce USP17 truncation expression. Three time points for expression of each truncation was tested: 4, 6, and 18 hours all at 25 °C with shaking at 200 rpm. The temperature and shaking speed was held identical to that of the full–length enzyme. Samples were lysed by OTG detergent, as described below, and analyzed by SDS–PAGE. Regardless of the time allowed for expression, large quantities of USP17 was clearly visible in the inclusion body indicating that even the truncations were still insoluble when expressed within *E. coli*. Refolding attempts were made with the catalytic truncation, residues A80–E390, identically to wild–type USP17 described above, however the protein precipitated upon refolding. No further attempts were made to refold the other truncations.

2.3.7 Expression and Purification of USP17 from the Baculovirus Expression System

Many full–length USPs are successfully expressed from the Baculovirus Expression System (BES) to optimize for solubility and properly folded enzymes. The BES used for the expression of USP17 was the BD BaculoGold system [50]. This system utilizes the *Autographa californica* nuclear polyhedrosis virus (AcNPV) for the recombinant expression of USP17. The *usp17* gene encoded in the modified pVL1392 vector (described above) was introduced into the AcNPV vector through homologous recombination and replaces the viral gene for polyhedrin, a gene not essential for the tissue culture lifecycle of the virus [50]. The recombination event requires the co–transfection of linearized AcNPV DNA and the *usp17*–pVL1392 into the *Spodoptera frugiperda, (Sf9)* insect cells. In general, recombinant proteins produced within this system can range between 0.1 – 50% of the total insect cell proteins produced [50]. The materials utilized for the recombinant expression of USP17 within BES are described within Table 2.6.

Maintaining *Sf9* Cells

Sf9 cells were maintained in suspension in 250 mL flasks with 20 mL aliquots at a density ranging between 80–350 x 10^4 cells/mL at 28 °C with shaking at 90 rpm. Once cells reached the middle of the log phase of their growth, a density of 250–350 x 10^4 cells/mL, they are split to a density of 80–100 x 10^4 cells/mL in supplemented media (*Sf9*–II SFM (1x), 5% FBS, and 1x Antibiotic–Antimycotic). Once the cell passage number reached 50, cells were disposed of and a new passage was thawed. Viability of cultured cells was analyzed every two days by hemocytometer.

Table 2.6
Baculovirus Expression System Materials

Material	**Vendor**	**Catalog Number**
Linearized Baculovirus DNA	BD BioSciences	Discontinued
USP17–pVL–1392	Synthesized by BioBasic	N/A
SF9– II SFM (1x) Media	Life Technologies	10902–088
100x Antibioitic–Antimycotic	Life Technologies	15240–062
Fetal Bovine Serum (FBS)	Life Technologies	16140–071
Cellfectin II Reagent	Life Technologies	10362–100
Sf9	Gibco–Invitrogen	11496–015
5 Cellstar Greiner	Sigma–Aldrich	658170
T175 Flasks Nunclon Delta Surface	Thermo Scientific	178883
Falcon Multiwell 6–Well	Falcon	353646

Freezing *Sf9* Cells

Sf9 cells to be frozen were grown to a density of 2–3 x 10^6 cells/mL and centrifuged at 500 x g for five minutes at room temperature in a 50 mL Falcon Tube. The media was aspirated, and the cells were resuspended in supplemented media to a density of 1 x 10^7 cells/mL. Resuspended cells were aliquoted in 1 mL volumes into cryogenic vials. The vials were pre–labeled with the date, cell–line, and passage number. The vials were placed into a cryosafe cooler filled with isopropanoyl and stored at –80 °C overnight. The next morning the cells were transferred to a liquid nitrogen cell dewer in the vapor phase (not submerged in liquid nitrogen).

Thawing *Sf9* Cells

Sf9 cells were removed from the vapor phase of the liquid nitrogen cell dewer and were rapidly thawed in a water bath at 37 °C for two minutes. Thawed cells were

then gently transferred to 30 mL of freshly prepared room temperature, supplemented media. Cells were centrifuged at 1000 x g for five minutes and the media was aspirated. Cells were resuspended in 20 mL of fresh supplemented media and grown in suspension a 250 mL flask.

Transfection, Homologus recombination of AcNPV Baculovirus DNA with *usp17*, and P1 Virus Production

A six–well plate was seeded to 70–80% confluency with *Sf9* cells, $\sim$ 2.6 10^6 cells per well and were allowed to adhere for five minutes at room temperature. The media was aspirated and 1 mL of unsupplemented *Sf9*–II SFM (1x) media was added to each well. The cells were incubated for one hour at 28 °C. During the incubation period, the mixtures displayed in Table 2.7 were made for USP17, but also for any control vectors (i.e. Empty Vector pVL1392 and USP7–pVL1392) were prepared as follows. Mix A was prepared in a 1 mL eppendorf tube as described in Table 2.7 and incubated for 30 minutes. Twenty minutes into Mix A's incubation, Mix B was prepared as described in Table 2.7. Mix B was allowed to incubate for five minutes before 100 μL of unsupplemented media was added, and mixed by pipetting up and down once, SLOWLY. Mix A was then added to Mix B, and mixed by slowly swirling. The combined mixture was allowed to incubate for 30 minutes. After the one hour incubation period the cells were evaluated, and the wells with the best confluency were chosen for transfection. The cells were transfected by adding the mixture drop–wise to the appropriate wells. The plate was swirled to combine and allowed to incubate at 28 °C for four hours. The cells were then evaluated to confirm they were still fixed to the plate and alive. At that time the media was aspirated, and 2 mL of fresh supplemented media was added to each transfected well. The cells were then incubated at 28 °C for 5 days to grow a P1 virus stock.

Table 2.7
Baculovirus Transfection Mixes

Mix A	**Mix B**
8 μL Cell–fectin	1.5 μL USP17–pEV–L8 Vector (1–2 ng/μL)
100 μL *Sf9*–II SFM (1x) Media	1.5 μL Linearized Baculovirus DNA
– Gently mix by pipetting	–Stir with pipette to mix

P2 Virus Production

Two T75 flask per P1 viral stock generated was seeded with 5 mL of *Sf9* cells at a density of 2–3 x 10^6 cells/mL and were allowed to incubate for five minutes at room temperature within the fume hood. The flasks were then evaluated for 70–80% confluency and were placed at 28 °C for at least one hour to allow the cells to adhere. In the case of successful homologous recombination for both *usp17* and the empty vector control P1 viral stocks can be evaluated for fluoresence of the GFP tag. If weak fluorescence was observed the P1 viral stocks were harvested. After the one hour incubation cell attachment was confirmed, and the media was aspirated. The P1 viral stock of USP17 was added drop–wise to one flask while the empty vector control P1 stock was added drop–wise to the other. The flasks incubated for one hour at 28 °C with gentle rocking back and forth every 20 minutes to ensure even coverage of the cells with virus. After the one hour incubation was complete, 10 mL of supplemented media was added to the flask and the flasks were placed at 28 °C for two days.

P3 Virus Production

The T75 flasks were evaluated for fluorescence of GFP to confirm successful P2 viral production. If P2 virus could be detected, the media was harvested containing the P2 viral stock. Two T175 flasks were seeded with 15 mL of *Sf9* cells at a density

of 2–3 x 10^6 cells/mL with an additional 5 mL of supplemented media. Again, one flask was for *usp17* and the other for the empty vector control. The cells were allowed to incubate for five minutes, and 70–80% confluency was confirmed. The cells were then stored at 28 °C for one hour to allow them to adhere. After the incubation, the media was aspirated from the flasks and 2 mL of the P2 viral stocks was added drop–wise to each the flask accordingly. The flasks were stored at 28 °C for one hour, rocking every 20 minutes to ensure even coverage of the cells with virus. After one hour, an additional 20 mL of supplemented media was gently added to each flask, and the flasks were placed at 28 °C for three days.

USP17 Protein Expression from *Sf9* cells

The T175 flasks were evaluated for fluoresence of GFP to confirm successful P2 viral production after 3 days of expression. If fluoresence levels were low, cells would be incubated for additional time up to five days total from the initial infection day. The P3 virus was harvested, and stored in a 50 mL Falcon tube, parafilmed, at 4 °C for up to six months. At least two days prior to protein expression, the cells were split to a density of 1–2 x 10^6 cells/mL in a volume of 200 mL in a 2 L Fernbach flask. The day before expression the density of the cells was evaluated, and the cells were again split to 1–2 x 10^6 cells/mL in 200 mL in 2 Fernbach flasks, one for USP17 expression and one for empty vector expression (i.e. GFP only). The morning of expression, the density of the cells was again evaluated to confirm the cells were at a density of 2.5–3.5 x 10^6 cells/mL. If the density was low, the cells were allowed to grow for an additional day to ensure the proper density at time of infection. To infect 200 mL of cells at a density of 2.5–3.5 x 10^6 cells/mL for protein expression, 8 mL of P3 viral stock was added. The flasks were allowed to shake a 90 rpm at 28 °C for three days.

The cells were harvested by centrifugation in a Sorvall RC 6+ ThermoScientific Centrifuge with a Fiberlite F9–4x1000y rotor at 3000 x g for 20 minutes at 4 °C. The cells were re–suspended in a buffer containing 500 mM NaCl, 10 mM Imidizole, 10

mM βME, and 20 mM Tris pH 7.5 and lysed by sonication with a Branson Digital Sonifer at 50% with pulsing of five seconds on and five seconds off for a total of four minutes. The lysate was clarified by centrifugation with a Beckman Coulter Avanti J–30I centrifuge in the JA–17 rotor at 80,000 x g for one hour at 4 °C. The clarified lysate was initially analyzed for USP17 expression via SDS–PAGE and fluorescence imaging (Figure 2.4). The resulting SDS–PAGE gel was placed on an UV lightbox to observe the GFP tag. The gel was then stained with Coomassie Brilliant blue for visualization of all the protein bands. The presence of multiple bands in the SDS–PAGE imaged by UV indicates that USP17 is actively degraded by proteases within the *Sf9* cells. As significant degradation and low yields of expression ($\leq$ 0.5 mg/ 200mL culture) of USP17 within this system was observed, this method was not further pursued for large scale recombinant USP17 production.

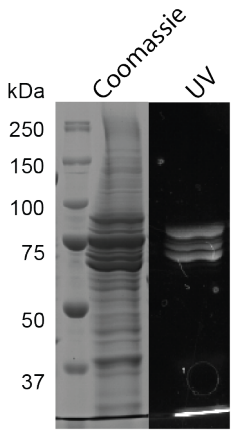


Fig. 2.4. **Purification of USP17 from Baculovirus Expression System.** *Left*–SDS–PAGE evaluation of the *Sf9* cell lysate of USP17 expression from the Baculovirus Expression System. *Right*–SDS–PAGE evaluation by UV imaging of the *Sf9* cell lysate of USP17 expression from the Baculovirus Expression System.

2.3.8 Measuring Protein Concentration

Microassay Bio–Rad Bradford Protein Assay

This assay was used throughout purifications to estimate the total protein concentration in the sample. The coomassie dye was prepared by diluting the Bio–Rad 5 x concentrate to 1x with nanopure water. A standard curve was prepared using Bio–Rad BSA powder that had been diluted to a stock concentration of 1.36 mg/mL in nanopure water. This stock was then diluted to a working stock of 0.5 mg/mL in the same purification buffer the protein sample to be measured was currently residing. This dilution was made with 18.4 μL of the BSA stock diluted into 31.6 μL of purification buffer. A standard curve of 0, 0.125, 0.25, 0.375, and 0.5, mg/mL was built by pipetting the following volumes of the working stock of BSA into a 96–well clear, flat–bottom microtiter plate: 0, 2.5, 5.0, 7.5, and 10 μL, respectively. The protein samples were added to the plate in a volume of 1, 5, or 10 μL depending on how concentrated the sample to allow the staining to fall within the standard curve. To each well 200 μL of the 1x Bio–Rad coomassie dye was added to each well containing protein or the BSA standard curve. Using a microplate mixer, the plate was mixed thoroughly for 30 seconds, and any bubbles within the wells were popped. Using the BioTek H1 Synergy plate reader, the absorbance of each well was read at a λ of 595nm. The absorbance of the BSA standard curve was plotted in terms of the corresponding concentration and fit using linear regression. The resulting standard curve was then used to estimate the protein concentration in the protein samples.

A_{280} Determination of Protein Concentration

To obtain a more accurate measurement of pure protein the concentration was measured using absorbance, rather than the Microassay Bio–Rad Bradford Protein Assay. The A_{280} of pure samples was measured using the Take3 microplate in 2 μL volumes with a pathlength of 0.05 cm. The protein concentration was determined

using the Beer–Lambert equation with the experimentally determined extinction coefficient for that sample.

2.3.9 Determination of the Extinction Coefficient for USP17

The concentration of purified USP17 and USP17C89S was determined using the Microassay Bio-Rad Bradford Protein Assay. Each protein was diluted to 1 mg/mL in both Size Exclusion Buffer (native state) and 6 M GuHCl (denatured state). The extinction coefficient for purified USP17 was experimentally determined as described by von Hippel et al [51]. The Take 3 microplate was used to measure the A_{280} for the protein diluted in buffer and GuHCl, each in triplicate, with the BioTek H1 Synergy. The background absorbance was subtracted from the measured absorbance for the protein sample. Using the Beer–Lambert equation, the molar extinction coefficients were determined for both the native and denatured proteins. The true molar extinction coefficient was then determined by dividing the native molar extinction coefficient by the denatured molar extinction coefficient, and multiplying by the computationally estimated molar extinction coefficient calculated from the primary sequence of the protein using Expacy's Protparam tool [52]. The molar extinction coefficient of USP17 was calculated as 58,104 $M^{-1}cm^{-1}$ and 58,338 $M^{-1}cm^{-1}$ for USP17C89S as described by von Hippel et al [51].

2.3.10 SDS–PAGE Analysis

Protein sample concentration was measured by a Microassay Bio–Rad Bradford Protein Assay. SDS–PAGE samples were prepared by mixing 5–10 μg of protein sample with 4 μL of 5x SDS Sample Loading Buffer (0.2 M Tris–HCl pH 6.8, 10% SDS, 10% glycerol, 0.02% Bromophenolblue, and 5 mM βME). Samples were heated at 95 °C for 2 minutes immediately prior to loading into the SDS–PAGE gel. Gels were run at 200 volts for 40 minutes. USP17 analyzed by 10% or 12.5% SDS–PAGE gel depending on the purpose.

2.3.11 OTG Lysis for Protein Expression Analysis

For each sample to be tested, 1.5 mL of the culture was micro–centrifuged in an eppendorf tube for 5 minutes at 13,500 x g at 4 °C. The supernatant was decanted and the cell pellet was re–suspended in 175 μL of OTG buffer (1% Octylthioglucoside (OTG) and 10 mM Tris pH 7.5). To the re–suspended sample, 0.01 mg/mL of DNAseI was added and gently vortexed to mix. The sample was centrifuged as described previously. The supernatant, corresponding to the soluble lysate, was removed, and stored on ice. The remaining pellet was re–suspended in 200 μL of OTG buffer supplemented with 5 μL of 10 mg/mL Lysozyme and vortexed for one minute. To the sample, 800 μL of of nanopure water was added the sample was vortexed again followed by centrifugation as described previously. The resulting supernatant was decanted and the pellet was washed with 1 mL of nanopure water and pelleted again. This was repeated once more to ensure the sample contains no lysozyme. The resulting sample is the inclusion body, or insoluble lysate, and was re–solubilized in 6 M GuHCl. The soluble and insoluble lysates were analyzed by SDS–PAGE.

2.3.12 Kinetic Assays of USP17

Specific Activity Assays

The specific activity was measured throughout each USP17 purification using the fluorogenic synthetic substrate ubiquitin–rhodamine 110 (Ub–Rho110, Boston Biochem). Ub–Rho110 is internally quenched when the rhodamine 110 molecule is conjugated to the ubiquitin moiety through a peptide bond. Upon cleavage of the peptide bond the quenching is released, and fluorescence from the rhodamine 110 is emitted. This substrate works as a generic substrate for many USP enzymes (Figure 2.5). The Ub–Rho110 is provided from Boston Biochem at a stock concentration of 250 μM in 100% DMSO which is normalized to 1 μM in Substrate Buffer (50 mM Tris pH 7.5, 1 mM EDTA, 100 mM NaCl, and 0.05% CHAPS) and diluted into each

assay at a final concentration of 500 nM. The protein concentration was measured by the Microassay Bio–Rad Bradford Protein Assay and was normalized in Assay Buffer (50 mM Tris pH 7.5, 0.1 mg/mL BSA, 5 mM DTT, and 0.01% $TritonX^{TM} - 100$) to a working concentration of 10 nM. The final concentration of protein in each activity analysis was 5 nM. First, 15 μL of the working concentration of Ub–Rho110 (1 μM) was added to a Costar 96 half–volume black plate, and the background fluorescence was read as a filter based assay using a BioTeK Synergy H1 plate reader every 10 seconds at an excitation λ of 485 nm and emission λ of 528 nm for two minutes. Following the background read, 15 μL of protein was added to the plate, the plate was mixed for 10 seconds and the assay was read under the same conditions for 30 minutes total or until the reactions started to plateau.

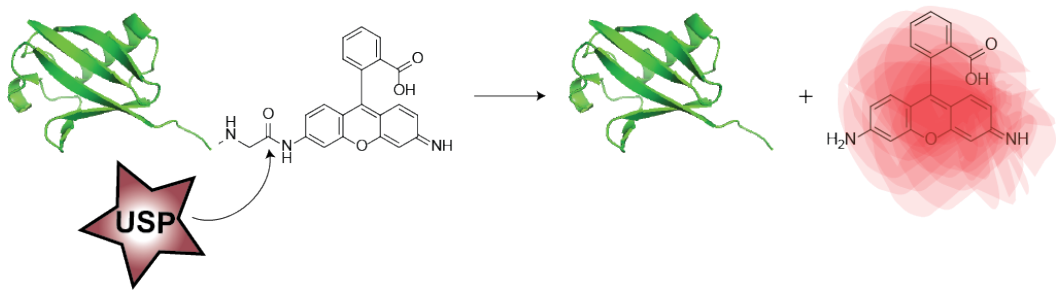


Fig. 2.5. **Fluorogenic Kinetic Assay for USPs.** The fluorogenic substrate Ub–Rho (as well as Ub–AMC, not shown) is composed of a ubiquitin molecule conjugated through a peptide bond to a Di–glycl Rhodamine 110 molecule. Many USPs are able to cleave the peptide bond similarly to what is described in Figure 1.4. Once the peptide bond between the ubiquitin and rhodamine molecules is broken fluoresence is emitted and is measured.

Determination of the Extinction Coefficient of the Substrate

The extinction coefficient of Ub–Rho110 was determined anew each day of the purification. To determine the extinction coefficient, five serial dilutions of Ub–Rho110 were made to the working concentrations of 0.03 μM to 1 μM. First, 15 μL of each

dilution of Ub–Rho110 was added to the plate. Second, 15 μL of a highly concentrated USP stock was added the plate and mixed. The plate was read identically to that of the specific activity assays. The reactions were allowed to plateau and the maximum artificial fluorescence units (AFU) was measured. The resulting maximum AFU values were plotted in terms of the corresponding concentration of Ub–Rho110. The plot was fit using linear regression and the resulting slope corresponds to the extinction coefficient in units of AFUs per concentration of Ub–Rho110 (AFU/μM).

Specific Activity Analysis

The initial slope of each specific activity reaction in units of fluorescence per unit time (AFU/min) was converted to the amount of product produced per unit time (μM/min) with the determined extinction coefficient of the substrate. Units, as defined by μmol/min, were obtained for each sample by dividing by the total assay volume. The specific activity of the sample, Units/mg of enzyme, was determined by dividing by the milligrams of protein present in the assay. The fractional recovery of activity was determined by dividing the total units of the sample by the total units of the refolded protein then multiplying by 100 to obtain a percentage. The fold in purification after each chromatography step was determined by dividing the specific activity of each sample by the specific activity of the refolded protein.

Steady–State Kinetic Assays

A fluorogenic activity assay for USP17 was developed in a 96–well, black half–volume plate (Corning Costar) to evaluated the kinetics of the deubiquitinating activity of USP17. Two substrates were used, ubiquitin 7–amino–4–methylcoumarin (Ub-AMC, LifeSensors) and Ub–Rho110. Assays were performed in 30 μL reaction volumes and in triplicate. Enzymatic reactions were initiated by the addition of 15 μLof enzyme followed by brief shaking (10 seconds) using the plate reader shaker. The rate of hydrolysis was analyzed at 10 second intervals over a time period of 30 minutes

using a BioTeK Synergy H1 plate reader equilibrated at 25 °C. Both substrates were measured as a filter based assay. The Ub–Rho110 was measured at the wavelengths described previously, while Ub–AMC was measured at an excitation λ of 360 nm and emission λ of 460 nm. Background fluorescence was measured first by pipetting 15 μL of the serially diluted substrate into the plate and at the appropriate wavelengths for 2 minutes total as described previously. Following the background read, 15 μL of the working stock of USP17 was added to the plate, the plate was mixed for 10 seconds, and the assay was read under the same conditions for 30 minutes total, or until the reactions plateaued. The rate of hydrolysis was analyzed in 10 second intervals for 30 minutes using the BioTeK Synergy H1 plate reader equilibrated at 25 °C. The initial slope of each reaction in Arbitrary Fluorescence Units (AFU) per unit time (AFU/min). These values were then converted to the amount of product produced per unit time (μM/min) using the fluorescence extinction coefficient of the product (either released AMC or Rho110). The extinction coefficient of both substrates was determined as described previously described for Ub–Rho110 in Section 2.3.12.

To determine the K_m and the k_{cat} values of USP17, the enzyme concentration was held constant at 3.125 nM and the substrate concentrations (Ub–AMC or Ub–Rho110) were varied from 0.01 μM to 10 μM. The enzyme was diluted in Assay Buffer, and the substrates were diluted in Substrate Buffer. The resulting rate values in μM/min were then converted into turnover number in units of sec^{-1} by dividing the rates by the enzyme concentration (V/[E]). These apparent turnover values were then plotted as a function of substrate concentration, and the data were fit to the Michealis–Menten equation (Equation (2.1)) using non–linear regression and the Enzyme Kinetics Module in the program SigmaPlot (v12: Systat Software Inc.).

$$v = \frac{V_{max}[S]}{K_m + [S]} \tag{2.1}$$

2.3.13 Evaluation of USP17 Stability

Aliquots of purified enzyme at a concentration of 1.26 mg/mL were incubated at 4 °C, 25 °C, or 37 °C for 1 to 3 days. At each time point, the sample was centrifuged at 13,500 x g for five minutes and the protein concentration was measured by A_{280} using the determined extinction coefficient of USP17 (Section 2.3.9). The specific activity of the samples was determined as described above except with a final concentration of enzyme of 6.25 nM of the enzyme and 0.5 μM of the Ub–Rho110. The percent remaining activity was determined at the same enzyme and substrate concentration by calculating V/[E] as described in Section 2.3.12, and calculating a percentage from the day 0 V/[E] measurement.

2.3.14 Analytical Size Exclusion Chromatography

The proteins ferritin, aldolase, conalbumin, and ovalbumin from GE Healthcare HMW calibration kit (Cat# 28–4038–42) was diluted to 3 mg/mL in Analytical Buffer(50 mM HEPES pH 7.5, 700 mM GuHCl, and 2 mM Tris(2–carboxyethyl) phosphine (TCEP)), with the exception of ferritin which was diluted to 0.3 mg/mL. BSA was diluted to 1.23 mg/mL. The proteins were combined into three separate injections to optimize for resolution: aldolase and ovalbumin, conalbumin and ferritin, and BSA. Each set of proteins was injected onto a 24 mL Superdex 200 Increase 10/300 (GE Healthcare) 24 mL (V_t) analytical size exclusion chromatography (SEC) column using an isocratic gradient at 4 °C and a flow rate of 0.5 mL per minute. For each protein the elution volume (V_e) was measured. A dextran sample (included with the GE Healthcare HMW kit) was injected at 1 mg/mL to determine the void volume (V_o) of the column under these conditions. The average distribution coefficient ($K_{average}$) [53] for each protein was determined by

$$K_{average} = \frac{(V_e - V_o)}{(V_t - V_o)} \tag{2.2}$$

A standard curve of $\log(M_r)$ vs $K_{average}$ was plotted and fit using linear regression,

$$K_{average} = -0.2641(M_r) + 1.641 \tag{2.3}$$

An aliquot of 500 μL of USP17 at a concentration of 1.26 mg/mL in Analytical Buffer was injected onto the same column. The $K_{average}$ was calculated with Equation (2.2) and molecular weight was determined with Equation (2.3).

2.3.15 Analytical Ultracentrifugation

Analytical Ultracentrifugation was performed with the assistance of the Purdue Biophysical Analysis Laboratory. To determine the oligomeric state of USP17, sedimentation velocity experiments were performed at 25 °C on the Beckman–Coulter XLA ultracentrifuge using 5.6 μM (0.4 mg/mL) USP17 in Analytical Buffer. The sample was run in an AN–60 Ti rotor at 50,000 rpm. Absorbance optics at 280 nm was utilized for protein detection. Solvent density, viscosity, and partial specific volumes were calculated using SEDNTERP. SEDPHAT was used to fit the data to estimate the sedimentation coefficient (s) to be 2.955 ± 0.002 which corresponds to a molecular weight estimate of 65.2 kDa.

2.3.16 Dynamic Light Scattering

Three milligrams of thawed USP17 was concentrated to ≤ 500 μL using ultrafiltration with a 10,000 MWCO membrane (Millipore) and was then injected onto a 24 mL Superose 6 Increase (GE Healthcare) SEC column equilibrated with Analytical Buffer. The column was run at a flow rate of 1 mL/min throughout. The eluted peak was collected as one fraction and analyzed by SDS–PAGE and the specific activity was determined. The fraction was then concentrated to above 1 mL and diluted to three concentrations, 0.5, 0.75, and 1 mg/mL, in Analytical Buffer. Dynamic light scattering (DLS) was performed for each USP17 concentration using a Malvern Zetasizer Nano S with a HeNe laser at 633 nm. The three concentrations of USP17

were measured at a fixed scattering angle of 173° (back scatter) in 100 μL aliquots at a constant temperature of either 4 °C or 25 °C. Size analysis was performed by number distribution curves, and were utilized to determine the molecular weight and oligomeric state using the Malvern DTS Software. The number size distribution represents the number of particles in the different size bins as determined by the Malvern DTS Software.

2.3.17 Evaluation of pH on the Activity of USP17

The effect of pH on the activity of USP17 was tested from pH 6.0 to 10.5 in half–step increments using a wide pH range buffer system (75 mM Tris, 25 mM Acetic Acid, 25 mM MES, 25 mM Glycine, 0.1 mg/mL BSA, 0.01% $Triton^{TM}X-100$, and 5 mM DTT) at each pH tested [17]. USP17 was diluted to a final concentration of 6.25 nM and Ub–Rho110 was diluted in the same buffer at each pH to two concentrations (0.5 μM and 5 μM) that are below and above the K_m value. Enzymatic activity was measured in triplicate as previously described (Section 2.3.12). The resulting data were fit to a bell–shaped profile representing two ionizations, pK_a and pK_b, with the kinetics module of SigmaPlot (v13: Systat Software Inc.).

2.3.18 Ubiquitin Chain Processing Assay

The substrate specificity of USP17 was tested against the di–ubiquitin chains of the eight linkage types (Lys^6, Lys^{11}, Lys^{27}, Lys^{29}, Lys^{33}, Lys^{48}, Lys^{63}, and linear (Boston Biochem)), as well as Lys^{48}, Lys^{63} tetra–ubiquitin chains (LifeSensors), and Lys^{11} tetra–ubiquitin chains (Boston Biochem). For reactions containing USP17 a concentration of 40 nM was used. Each di–ubiquitin reaction contained 4 μM of the ubiquitin chain, and was incubated with or without enzyme at 25 °C for 120 minutes. The tetra–ubiquitin reactions contained 2 μM of the ubiquitin chains, and were incubated with or without enzyme for a range of times from 10 to 120

minutes at 25 °C. Reactions were analyzed by NuPAGE Novex 4-12% Bis-Tris mini gels (Invitrogen) stained with Coomassie Brilliant Blue.

2.4 Results

The methodology for the expression and purification of recombinant USPs is diverse. A number of expression constructs and expression methods have been used to produce recombinant USPs, and a summary of these approaches is provided in Table 2.8. While there are a few USPs (i.e. USP12 and USP16) that can be expressed and purified from *E. coli* as full–length constructs, many USPs are insoluble when expressed from *E. coli* [11, 54]. An alternative approach for the expression of full–length USPs has been to utilize the labor–intensive baculovirus expression system (BES) which has successfully produced full–length USP1, USP2, and USP7 proteins [11, 17, 54]. However, the most common approach for producing recombinant USPs is to truncate the USP of interest into its individual domains in order to obtain high yields of soluble protein within *E. coli*. With this approach, high priority is placed on the catalytic domain, as demonstrated for the well–characterized USP7, USP8, and USP14 catalytic domain constructs [11, 16, 18–20, 54]. However, it is often found that the enzymatic activity of the catalytic domain is significantly differs from that of the full–length enzyme, making production of full–length USPs a necessity for advanced studies. We therefore pursued approaches for the production of full–length USP17.

2.4.1 High yields of USP17 can be obtained by refolding from bacterial inclusion bodies

USP17 with a N–terminal 6His tag was expressed and purified from *E. coli* inclusion bodies. The expression and refolding procedure involved washing the resulting inclusion bodies with 1% $Triton^{TM}X-100$ by homogenization, and then solubilizing the protein using 6 M GuHCl. Refolding of active USP17 was achieved by rapid dilution of the protein into a refolding buffer containing only 0.7 M GuHCl. Purification

Table 2.8
USP Expression Systems

	Construct	Expression System	Reference
$USP1$	Full–Length	BES	Villami et al. [17]
$USP7$	Full–Length	BES	Faesen et al. [20]
$USP7_{CD}$	Catalytic Domain	*E. coli*	Faesen et al. [20]
$USP8_{CD}$	Catalytic Domain	*E. coli*	Avvakumov et al. [18]
$USP11$	Full–Length	*E. coli*	Faesen et al. [11]
$USP12$	Full–Length	BES	Luna–Vargas et al. [54]
$USP14$	Catalytic Domain	*E. coli*	Hu et al. [16]
$USP16$	Full–Length	*E. coli*	Luna–Vargas et al. [54]
$USP17$	Full–Length	Refolded	This work

of the active, refolded USP17 was achieved in two steps including nickel–metal–chelate affinity chromatography followed by SEC. From 5 L of *E. coli* culture, approximately 21 mg of USP17 can be obtained and the enzyme is >95% pure as judged by SDS-PAGE (Figure 2.2). From SDS-PAGE analysis, USP17 has an estimated weight of 60 kDa, that is close to the expected size of USP17 with a 6His tag (∼62 kDa). The specific activity and yield of USP17 was measured at each step during the purification, and the results are summarized in Table 2.5. The activity measurements are normalized to the refolding step, the first step where fully active enzyme is present. The resulting calculated yield of USP17 was 10.5%, and the overall improvement in purification was 5–fold.

2.4.2 USP17 degradation during expression

During the development of the USP17 purification procedure, it was evident from SDS–PAGE analyses and confirmed by western blot with an antibody towards the

6His tag that USP17 is prone to some truncation. These truncations are either the result of self–degradation or USP17 or the susceptibility of USP17 to protease degradation during expression (Figure 2.2A). While a high yield of the full–length enzyme can be purified in high yield away from the truncations by the refolding method described above (Figure 2.2A), three different expression methods were attempted to identify the cause and reduce the amount of degradation observed. First, USP17 was expressed with a N–terminal 10His–GFP tag by baculovirus infection of *Sf9* cells. The yield of USP17 from *Sf9* cells was significantly less than what can be obtained by refolding, and degradation was still apparent as observed with a GFP–tagged version of USP17 in Figure 2.4, lane 2. Second, we expressed, refolded from inclusion body, and purified a catalytic mutant of USP17, USP17C89S Figure 2.3, to evaluate if the degradation is autocatalytic or the result of proteases within the cell during expression. During the purification of USP17C89S the degradation pattern of USP17 was again observed Figure 2.3 indicating that USP17 is susceptible to protease degradation during expression. This conclusion is further supported by the absence of truncations appearing after the purification is complete, suggesting purified full–length USP17 is stable and not prone to self–degradation. Third, attempts were made to truncate USP17 to improve solubility and prevent degradation Figure 1.5, a technique utilized by many other USPs as summarized in Table 2.8. However none of the USP17 truncations were soluble or able to refold in a stable form from *E. coli* inclusion bodies. In summary, the refolding and purification methodology described in Section 2.4.1 is the best procedure for obtaining high yields of recombinant full–length USP17.

2.4.3 Important Refolding Buffer and Purification Observations

Three significant observations were made regarding the stability of USP17 during the optimization of the refolding and purification of USP17 from inclusion body. First, during the optimization of the refolding buffer, it was found that if the GuHCl concentration was reduced to < 0.7 M USP17 precipitated immediately upon rapid

dilution. Second, while USP17 can actively refold at a pH of 7.5, significantly more precipitant is observed than when USP17 is refolded at pH 9.5. Finally, to obtain a high yield of pure full–length USP17, 5% glycerol is required in the refolding buffer, and all subsequent purification buffers to aid in stabilization. Without the glycerol present, USP17 would precipitate throughout the purification, specifically when concentrated for SEC. Furthermore, the glycerol acts as a cryo–protectant when USP17 is flash frozen in liquid nitrogen and stored at -80 °C.

2.4.4 Refolded USP17 is unstable at temperatures $\geq$ 25 °C

The stability of USP17 at different temperatures was tested by incubating USP17 at 4 °C, 25 °C, or 37 °C for 1–3 days. After incubation, the samples were centrifuged and the remaining USP17 concentration, the remaining activity (Figure 2.6A), and the specific activity were measured (Figure 2.6B). Large amount of precipitation was visualized in the samples incubated at 37 °C for each day. Almost a complete loss in both activity and specific activity was measured on day one and each day after, suggesting the remaining soluble protein was inactive. Large amounts of precipitation was again visualized for samples stored at 25 °C. After day one of the 25 °C samples only ∼10% of the activity remained, a substantial loss that is reflected in the specific activity measurement as well. Minimal remaining activity or specific activity remained for days two and three for the samples stored at 25 °C suggesting the remaining soluble enzyme was inactive. The samples stored at 4 °C had minimal precipitation and only lost ∼35% of the total activity after day one. This trend was reflected in the specific activity measurement as well. Unlike the other temperatures tested, no further loss in total activity was observed after days two and three for the 4 °C samples, suggesting that any aggregation or inactivation that occurred during incubation happened within the first 24 hours and not thereafter. The remaining specific activity over time and temperature are summarized in Figure 2.6.

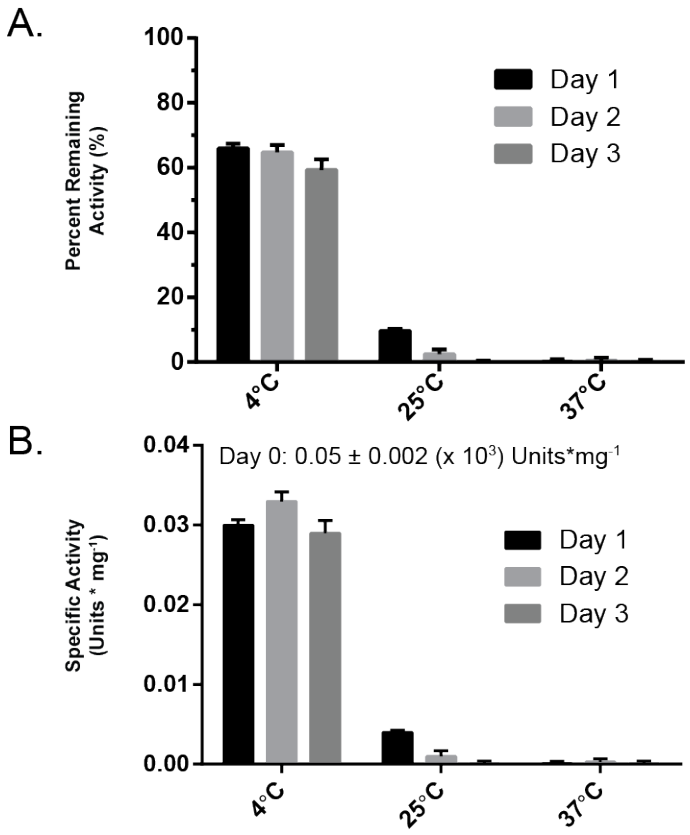


Fig. 2.6. **Retention of USP17 Specific Activity** Aliquots of USP17 were stored at 4 °C, 25 °C, or 37 °C for one, two, or three days. At each time point the sample was centrifuged and the remaining enzyme concentration was measured by A_{280}, and adjusted with the determined extinction coefficient to 6.25 nM. **A.** Remaining activity was measured with 0.5 μM Ub–Rho110. **B.** Specific activity was calculated as described in Section 2.3.13 with the same concentrations of substrate and enzyme as **A**. All assays were done in triplicate. Error, SD.

2.4.5 USP17 is a monomer in solution

To analyze the homogeneity and oligomeric state of refolded USP17, three methods were used: analytical–SEC, Analytical Ultracentrifugation (AUC), and dynamic light scattering (DLS). For each technique USP17 was diluted in the same Analytical Buffer for consistency.

Analytical Size–Exclusion Chromatography (SEC)

A 500 μL sample of USP17 at a concentration of 1.26 mg/mL was passed over a 24 mL Superdex 200 (GE Healthcare) SEC column at 4 °C. The elution profile (Figure 2.7A) displayed a broad peak, and the average distribution constant, $K_{average}$, of the maximal height and the half maximal heights was calculated. Using the standard curve constructed under the same conditions (Figure 2.7B), the molecular weight of the peak corresponds to 122 kDa and ranged from 86 to 153 kDa.

Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) was also performed to determine the oligomeric state of USP17 in solution. For this technique, 5.6 μM (0.4 mg/mL) sample of USP17 was analyzed at 50,000 rpm at 25 °C. Under these conditions, the resulting sedimentation coefficient was determined to be 2.955 ± 0.002 which corresponds to 65.2 kDa (Figure 2.7C).

Dynamic Light Scattering

The third method for analyzing the homogeneity and oligomeric state of refolded USP17 was DLS. USP17 was passed over a Superose 6 Increase (GE Healthcare) SEC column, and the peak fraction of pure USP17 was concentrated to 1 mg/mL. To evaluate the differences observed between the analytical–SEC and AUC results, USP17 was evaluated at 1, 0.75, and 0.5 mg/mL, and at two temperatures, 4 and 25 °C. The hydrodynamic radius measured at 4 °C (Figure 2.7D) ranged from 3.5 ± 0.8 nm and 3.75 ± 0.8 nm for 0.75 mg/mL and 0.5 mg/mL respectively. These numbers equate to a molecular weight range of 64 ± 2 kDa to 73 ± 2 kDa. The hydrodynamic radii broadened slightly for each concentration when measured at 25 °C (Figure 2.8). For example, at 0.75 mg/mL of USP17 the hydrodynamic radius was 3.5 nm at 4 °C and 4 nm at 25 °C, that increased the calculated molecular weight from 64 kDa to

93 kDa respectively. The hydrodynamic radius and the resulting molecular weights from all conditions tested are displayed in Table 2.9.

Table 2.9
Estimated molecular weight of USP17 by dynamic light scattering

		USP17 Concentration (mg/mL)		
Temperature		0.5	0.75	1
4 °C	R_H(nm)	3.7 ± 0.8	3.5 ± 0.8	3.6 ± 0.9
	MW (kDa)	73 ± 2	64 ± 2	67 ± 2
25 °C	R_H(nm)	3.6 ± 0.8	4 ± 1	4.1 ± 0.9
	MW (kDa)	67 ± 2	93 ± 3	90 ± 3

2.4.6 Steady–State Kinetic Characterization of USP17

The steady–state kinetic parameters were determined for USP17 by measuring the kinetic response of USP17 to increasing concentrations of two fluorogenic synthetic substrates, Ub–AMC and Ub–Rho110. As shown in Figure 2.9, USP17 can be readily saturated by both substrates. The data in Figure 2.9 were fit to the Michealis–Menten equation and the resulting kinetic parameters from those fits are summarized in Table 2.11. The kinetic parameters suggest that USP17 recognizes and hydrolyzes both Ub–AMC and Ub–Rho110 with nearly equal efficiencies (k_{cat}/K_m = 1500 $x10^3 M^{-1} sec^{-1}$ for Ub–AMC and 880 $x10^3 M^{-1} sec^{-1}$ for Ub–Rho110). Ub–AMC is the more commonly utilized fluorogenic substrate within the deubiquitinating enzyme literature. As such, Table 2.10 compares the kinetic parameters of USP17 with Ub–AMC as a substrate, to those of other human USPs characterized to date [11]. Interestingly, USP17 catalyzes the hydrolysis of Ub–AMC 3–fold more efficiently than USP7 as a result of the lower K_m value associated with USP17. Similar to other USPs, USP17 catalyzes poorly the hydrolysis of the fluorogenic peptide substrate RLRGG–

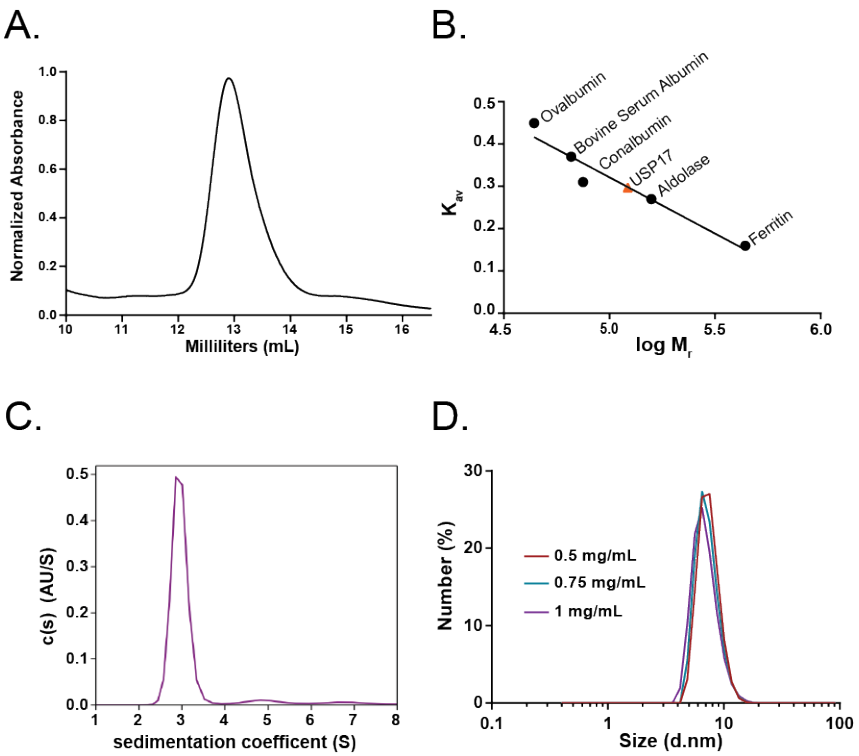


Fig. 2.7. **Determination of Oligomeric State of USP17. A.** Elution profile of USP17 from a 24 mL Superdex 200 (GE Healthcare) column. **B.** Standard curve built from the average distribution coefficient ($K_{avereage}$) vs log (M_r) of well characterized proteins, black circles. The $K_{average}$ of USP17 was determined from the elution profile in (A) and fit to the standard curve to determine the molecular weight of USP17, orange triangle. **C.** AUC sedimentation velocity (AUC-SV) analysis at loading a concentration of 5.6 μM of USP17. Plot of the distribution of sedimentation coefficients c(s) versus s, where s is plotted in Svedberg units, (S) calculated from AUC sedimentation velocity experiments. **D.** DLS curves at 4 °C of USP17 at three concentrations 0.5 mg/mL (deep red), 0.75 mg/mL (teal), and 1 mg/mL (purple). All assays were performed in triplicate.

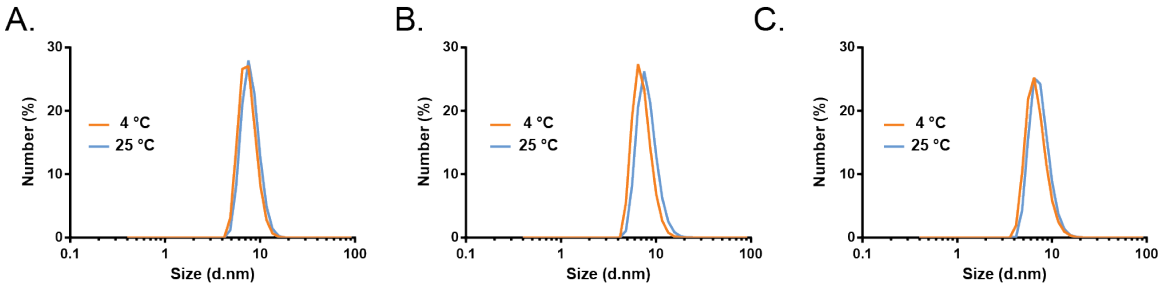


Fig. 2.8. **Summary of DLS Analysis of USP17 at Two Temperatures.** DLS curves of USP17 at two temperatures: 4 °C (Orange) and 25 °C (Blue) at three concentrations of USP17 **(A.)** 0.5 mg/mL, **(B.)** 0.75 mg/mL, and **(C.)** 1 mg/mL. All assays were performed in triplicate.

AMC and the ubiquitin–like modifier substrate ISG15–AMC, that is involved in the innate immune response.

Table 2.10
Kinetic Parameters of Well–Characterized USPs with Ub–AMC

	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($x10^{-3}s^{-1}M^{-1}$)
[a] $USP1_{21-785}$	9.71 ± 0.85	0.079 ± 0.003	9
[b] $USP2_{CD}$	2.4 ± 0.2	0.35 ± 0.03	146
[a] $USP7$	2.89 ± 0.1	1.37 ± 0.01	482
[a] $USP8_{CD}$	17.3 ± 2.5	7.90 ± 0.46	464
[a] $USP11$	0.77 ± 0.13	0.074 ± 0.003	0.147
[a] $USP12$	12.01 ± 4.2	0.0023 ± 0.002	0.2
[a] $USP16$	1.42 ± 0.25	0.069 ± 0.003	49.3
USP17	0.98 ± 0.06	1.47 ± 0.02	1500

[a] Values obtained from Faesen et al. [11]. [b] Values obtained from Bozza et al. [55]. USP17 values reported in this work.

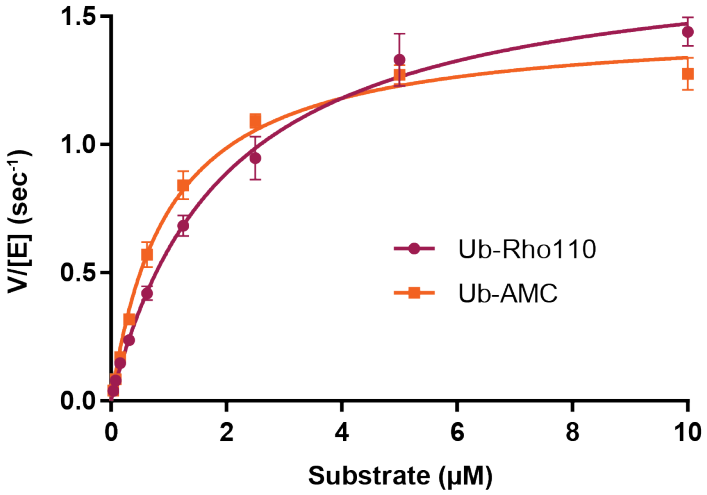


Fig. 2.9. **Kinetic response of USP17 to increasing concentrations of Ub–Rho110 and Ub–AMC** Michealis–Menten plots of USP17 with fluorogenic substrates: Ub–Rho110 (purple, closed circles) and Ub–AMC (orange, closed squares). Both substrates were evaluated in a concentration range of 0.04 to 10 μM with 3.125 nM USP17. Assays were performed in triplicate. Curves were fit to Equation (2.1), and the kinetic parameters, K_m and k_{cat}, were determined for each substrates.

Table 2.11
Kinetic Parameters of USP17 with Ub–AMC and Ub–Rho110

	Ub–AMC	Ub–Rho110
$K_m(\mu M)$	0.98 ± 0.06	2.0 ± 0.1
$k_{cat}(s^{-1})$	1.47 ± 0.02	1.76 ± 0.05
$k_{cat}/K_m(x10^{-3}s^{-1}M^{-1})$	1500 ± 90	880 ± 50

2.4.7 Effect of pH on USP17 Activity

The kinetic response of USP17 to variable pH values over the pH range of 6 to 10.5 was measured using Ub–Rho110 as a substrate and the results are shown in Figure 2.10. The rates of hydrolysis of Ub–Rho110 were evaluated at subsaturating (0.5 μM) and near saturating (5 μM) concentrations of Ub–Rho110 to look for an

observed shift in pK_a values. Ten pH values spaced 0.5 pH units apart were chosen, and the rates were measured in triplicate. The plots of the data in Figure 2.10 show typical bell–shaped curves, and therefore the data were fit to the equation that describes the kinetic model of an enzyme undergoing two ionizations with two pK_a values (pK_a and pK_b). The resulting curve fits are shown in Figure 2.10 for both subsaturating and saturating substrate concentrations, and the resulting pK_a for 0.5 μM Ub–Rho is 6.83 $\pm$ 0.07, which is similar to the pK_a value at 5 μM Ub–Rho110 which is 6.75 $\pm$ 0.04. The pK_b value at 0.5 μM Ub–Rho110 is 9.13 $\pm$ 0.07 which is similar to the pK_b value of 9.35 $\pm$ 0.04 at 5 μM Ub–Rho110. The lack of any significant differences in the pK_a and pK_b values at subsaturating and saturating concentrations suggests that ionizations in the free enzyme (E) and substrate–bound enzyme complex (ES) are the same, and that substrate binding does not influence the pK_a values.

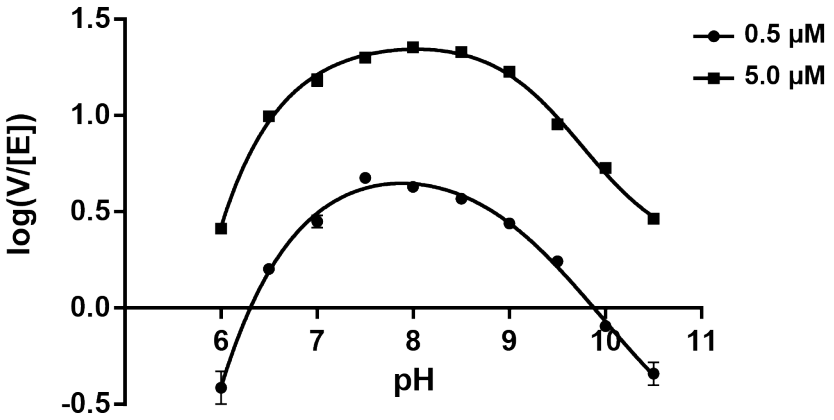


Fig. 2.10. **Kinetic response of USP17 to varying pH.** The activity of USP17 was measured with Ub–Rho110 from pH 6 to 10.5 with 6.25 μM USP17. Two concentrations of Ub–Rho110 surrounding the K_m value were tested, 0.5 μM and 5 μM. Assays were performed in triplicate. Curves were fit with a bell–shaped rate profile within SigmaPlot (v13: Systat Software Inc.).

2.4.8 Substrate Specificity of USP17 for different Ubiquitin Chain Linkages

Ubiquitination of protein substrates involves one of the seven lysine residues of ubiquitin. The topology of the chain is directly dependent on the specific lysine residue that is used to conjugate one ubiquitin to the next. These differences confer an additional layer of specificity for ubiquitin signaling beyond recognition of the protein substrate. Previous cell–based assays have demonstrated that USP17 can process Lys^{63} chains conjugated to RCE1 and SDS3 and Lys^{48} chains from CDC25A [29, 30, 47]. To expand upon this information, we performed a di–ubiquitin panel of all seven lysine linkages and the linear peptide linkage (linear) to determine the specificity of USP17. Di–ubiquitins with different isopeptide linkages (Lys^{6}, Lys^{11}, Lys^{29}, Lys^{33}, Lys^{48}, or Lys^{63}) or linear were incubated with USP17, and the reaction products were analyzed by SDS–PAGE Figure 2.11. USP17 was able to process all linkages to a certain extent, with the exception of linear. As was expected, USP17 was able to process Lys^{48} and Lys^{63} almost fully to mono–ubiquitin, as well as Lys^{11} and Lys^{33} (Figure 2.11 A, Top Panel). A substantial amount of uncleaved di–ubiquitin remained for Lys^{27}, Lys^{29}, and Lys^{6} chains (Figure 2.11 A, Bottom Panel). A comparison of these cleavage patterns to other characterized USPs is displayed in Table 2.12 [11]. This comparison suggests that USP17 is most similar to USP1, and is less promiscuous than other members of the USP family. Next we evaluated how USP17 processed ubiquitin chains by incubating USP17 with tetra–ubiquitin chains of either Lys^{48} (Figure 2.11 B), Lys^{63} (Figure 2.11 C), or Lys^{11} (Figure 2.11 D) by evaluating the cleavage pattern over time. The data suggests that all three chain types are processed by exo–trimming to a similar extent within the time range tested.

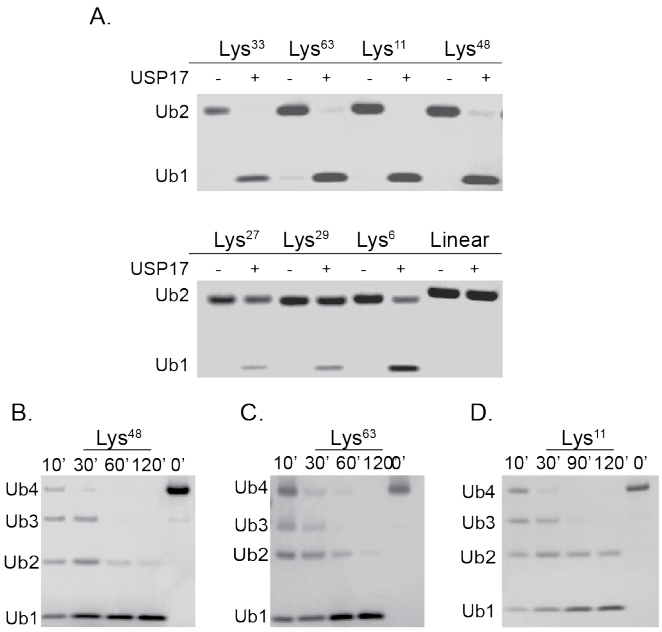


Fig. 2.11. **Ubiquitin Chain Recognition by USP17.** For all reactions, 40 nM USP17 was combined with ubiquitin substrate, and allowed to incubate at 25 °C. **A.** Survey of the hydrolysis of di–ubiquitin chains by USP17. Each assay contained 4 μM of the di–ubiquitin chain (Lys^6, Lys^{11}, Lys^{27}, Lys^{29}, Lys^{33}, Lys^{48}, Lys^{63}, and linear), and was allowed to incubate for 2 hr. Di–ubiquitin chains were incubated without enzyme as a negative control. **B, C, and D.** time–dependent hydrolysis of 2 μM Lys^{48} tetra–Ub (**B.**), Lys^{63} tetra–Ub (**C.**), and Lys^{11} tetra–Ub (**D.**) by USP17. Aliquots were removed at four time points and quenched with sample buffer, and analyzed by NuPAGE Novex 4–12% Bis-Tris mini gels (Invitrogen) stained with Coomassie Brilliant Blue.

Table 2.12
Comparison of USP17 Di–Ubiquitin Linkage Preferences of Various USPs

	Lys^6	Lys^{11}	Lys^{27}	Lys^{29}	Lys^{33}	Lys^{48}	Lys^{63}	*Linear*
[a]$USP1$	+++	++	+	+	++	+++	+++	-
[a]$USP7$	+++	+++	++	+	++	+++	+++	-
[a]$USP8_{CD}$	+++	+++	+++	+++	+++	+++	+++	-
[a]$USP11$	+++	+++	+++	++	+++	+++	+++	-
[a]$USP16$	++	+++	+++	+++	+++	+++	+++	+
$USP17$	++	+++	+	+	+++	+++	+++	-

[a] Values obtained from Faesen et al. [11]. USP17 values reported in this work.

2.5 Discussion

2.5.1 New methodology for expression and purification of USPs

The USP family of DUBs has a wide functional diversity in maintaining homeostasis within the cell through their regulation of many signaling pathways, including the ubiquitin–proteasome, DNA repair, and the cell–cycle. As a result, the mis–regulation of the USPs themselves frequently results in the development of cancer. With this realization, there is a great need to expand upon the number of USPs that have been expressed and purified allowing for more in depth kinetic, biochemical, biophysical, and structural studies of the USP family. Furthermore, these advanced approaches can be used in conjunction with the development of small molecule inhibitors to further probe the functional roles of the USPs in cells, and ultimately therapeutic compounds. Of the USPs that have been expressed via recombinant systems, few of them have yet to produce full–length enzymes due to their inherent insolubility when expressed in *E. coli*. Therefore, it is not uncommon to see full–length USPs expressed from the more labor–intensive, often low–yielding baculovirus expression systems. To obtain higher yields of soluble USPs, the individual catalytic domains or other domains of USPs have more often been produced.

Similar to other USPs, we found that USP17 is highly insoluble when expressed in *E. coli*. USP17 is a critical DUB involved in many pathways that dictate cell–cycle progression. Both *in cellulo* and *in vitro* based studies have elucidated the importance of USP17s activity on the stabilization of phosphatase CDC25A [29]. However, until our work reported here, recombinant expression and purification methods for USP17 had not been developed or reported, which greatly limited our ability to study this important enzyme. We were able to develop and optimize a reproduciable procedure for the expression, refolding, and purification of USP17 from *E. coli* inclusion bodies. With large quantities of recombinant USP17 in hand, we were able to interrogate the biochemical, kinetic and oligomeric properties of USP17, and determine its substrate specificity towards ubiquitin and ubiquitin–like substrates. The findings described

build upon our knowledge of the USP family and have important implications for the development of anti–cancer therapeutics.

2.5.2 Instability of USP17 may be a potential regulatory mechanism

Our observations during the expression of USP17 from both baculovirus *E. coli* reveled that USP17 is prone to proteolytic degradation. Native USP17 expression in human cells only occurs for a short period of time during the transitions from $G_1 - S$ and $G_2 - M$ checkpoints in the cell–cycle. When expression of USP17 occurs outside of these checkpoints, proliferation ceases. It is possible that this observed instability of USP17 is a regulatory mechanism within the cell, and as such, USP17 is prone to degradation to ensure that the presence of active USP17 is short lived, and that timely progression of the cell–cycle occurs. Alternatively, as is commonly seen with recombinant proteins, neither expression system may contain the required machinery to properly refold USP17, resulting in unfolded regions of USP17 being susceptible to proteolysis. It is not likely that USP17 itself is responsible for its degradation, as similar degradation patterns are observed for the refolded catalytic mutant. This evidence suggests that proteases native to *Sf9* and *E. coli* are responsible for the observed degradation of USP17.

2.5.3 Obscurity of the oligomeric state of USP17

Three approaches were utilized to determine the oligomeric state of USP17: analytical–SEC, AUC, and DLS techniques. USP17 was determined to be a monomer in solution by AUC when evaluated at 25 °C and a concentration of 0.4 mg/mL. It was also observed to be a monomer in solution at 4 °C and at concentrations ranging from 0.5 to 1 mg/mL by DLS. However, USP17 was determined to shift to a more dimeric state when evaluated by either analytical–SEC at 1.3 mg/mL and 4 °C or by DLS at 25 °C and at concentrations from 0.75 to 1 mg/mL. Taken together, the results suggest USP17 primarily assumes a monomeric state at lower concentrations, i.e. under the

conditions of our kinetic assays, but the equilibrium may shift towards a dimer at higher USP17 concentrations, i.e. <1 mg/mL, and this dimerzation may have some temperature dependence.

2.5.4 Kinetic efficiency of USP17 is significantly greater than that of other USPs

The kinetic parameters of USP17 were determined for both Ub–Rho110 and Ub–AMC substrates, and only modest differences in the kinetic parameters were observed. Of greater interest are the significant differences observed between the kinetic parameters for USP17 with Ub–AMC and the reported kinetic parameters for other USPs (Table 2.10) [11]. The catalytic efficiency USP17 ($k_{\mathrm{cat}}/\mathrm{K_m}$) is 3–fold higher than that of the very efficient and full–length USP7 [11], and is 7,500 times more efficient than the catalytic domain of USP2. A possible explanation for the higher observed catalytic efficiency for USP17 may reflect its minimal expression time within the cell–cycle. Without such high activity, USP17 may not be able to progress the cell–cycle efficiently within its short time frame or window of expression. A comparison to other USPs involved in proliferation supports this hypothesis. For example, the expression of USP1 is associated with low levels of DNA damage, and only has a short window of expression during the S–phase of the cell–cycle, similar to the expression pattern of USP17 [2]. However, USP1 is not required for cell–cycle progression, and, consequently, the activity of USP1 need not be as high as USP17 [2]. These differences may be further evidence of the significant role USP17 has in progressing the cell–cycle in a coordinated timely manner as compared to other USPs involved in cellular proliferation.

2.5.5 USP17 processes ubiquitin chains associated with degradation and enzyme activation

A number of USPs have so far been investigated for their ability to recognize and process the ubiquitin chains, and each has been found to exhibit its own degree of chain linkage specificty or promescuity (Table 2.12). This specificity is partially implied by the different topologies the specific ubiquitin chains assume and whether or not the USP is capable of accessing each iso–peptide bond. Faesen and coworkers performed a di–ubiquitin panel to assess the ability of a variety of USPs to process over time the eight different ubiquitin chain types [11]. While subtle differences could be observed between the USPs ability to process the chains, it was clear that the majority of the USPs tested were promiscuous, and could process the majority of the chains with the exception of linear chains [11]. USP17 appears to be a bit more specific about the topology of each linkage than other members of the USP family (Table 2.12). While USP17 could process each of the ubiquitin chain types, with the exception of linear ubiquitin, it did so to varying degrees. USP17 easily processed Lys^{48} and Lys^{63}, supporting the previous cell–based findings [29,30,47]. Interestingly, USP17 processed the Lys^{33} chain relatively well, a chain type that is not easily processed by most USPs, and has recently been associated with the modulation of kinase activity [8, 11]. USP17 also easily processed Lys^{11}, which has been described as a strong degradation signal, similar to the Lys^{48} chains [7, 11, 46].The observed cleavage pattern by USP17 is most similar to the cell–cycle dependent USP1, whose importance was described in Section 2.5.4.

We also determined the efficiency by which USP17 processes longer ubiquitin chains, including Lys^{48}, Lys^{63}, and Lys^{11} linked tetra–ubiquitin chains was examined. The cleavage pattern for both chain types suggest USP17 efficiently deubiquitinates each of these substrates, and utilizes an by exo–trimming mechanism, as no accumulation of a single intermediate ubiquitin species other than mono–ubiquitin was observed [12]. Furthermore, as these chains assume significantly different topologies,

these findings suggest that the specificity of USP17 is more likely dependent on the protein substrate the ubiquitin chain is conjugated to, not the chain itself [12].

2.5.6 USP17 activity is optimal between pH 7 and 9

The effects of pH on the ability of USP17 to process substrate was evaluated by varying the pH and Ub–Rho110 concentration. When the data was plotted, both concentrations of Ub–Rho110 surround the K_m (0.5 and 5 μM) in terms of pH assumed a bell–shaped curve. By analysis of these curves, the pK_a and pK_b values could be determined from the acidic limb (left hand side) and the basic limb (right hand side), respectively. The resulting pK_a for 0.5 μM Ub–Rho was 6.83 $\pm$ 0.07, which was similar to the pK_a at 5 μM Ub–Rho110 of 6.75 $\pm$ 0.04. This similarity was reflected in the pK_b values with 0.5 μM Ub–Rho110 at 9.13 $\pm$ 0.07 compared to 5 μM Ub–Rho110 at 9.35 $\pm$ 0.04. These values most likely correlate to the catalytic histidine (pK_a) and cysteine (pK_b) residues of the enzyme. The estimated pK_a values suggest that at pH values below 6.75 the catalytic histidine will exist predominantly as a positively–charged imidazolium ion, a state unable to deprotonate the cysteine to perform catalysis. The estimated pK_b values suggest that at pH values above 9.35 the catalytic cysteine will exist predomainatly as a thiolate and not require deprotonation. Therefore, only at pH values between 6.75 and 9.35 will both the cysteine and histidine exist in the proper ionization states to perform the following catalytic steps (Figure 1.4):

- The nucleophilic attack on the substrate to form the ES complex
- The general–acid assisted protonation of the leaving group in the FP state
- The deprotonation of the water molecule in the F state
- The protonation of the thiolate in the EQ state

Furthermore, at pH values above 9.35 the catalytic histidine residue will be deprotonated, therefore not capable of the general–acid assisted protonation of the leaving

group in the FP state. Minimal differences were observed between the estimated pK_a and pK_b values at the two substrate concentrations suggesting that the ionization state of the substrate is not influencing the formation of the USP17– Ub–Rho110 complex. Furthermore, as the difference in the pK_a and pK_b values at the two substrate concentrations is minimal, the pH does not have a large affect on the K_m or k_{cat} of the system.

CHAPTER 3. EXPRESSION AND PURIFICATION OF USP7 AND USP7 TRUNCATIONS

Parts of the data and text in this chapter have been published in various journal articles (insert references here).

3.1 Introduction

Human USP7 is a 125 kDa protein that is composed of seven different domains, an N–terminal Tumor Necrosis Factor receptor associated factor (TRAF) domain, the conserved papain–like catalytic domain, and 5 consecutive ubiquitin like (UBL) domains (Figure 3.1). The X–ray crystal structures have been solved for each of these domains by others in the field, however a structure of full–length human USP7 ($USP7_{FL}$) remains to be elucidated [19, 20, 33]. Our studies with USP7 have focused on elucidating the interactions between these different domains using steady–state kinetic, isothermal calorimetry, and gel–shift approaches. To perform these types of experiments eight different truncations of USP7 were cloned, expressed, and purified. This chapter describes the design, synthesis, sub–cloning, expression, and purification of each USP7 construct depicted in Figure 3.1 that were utilized in Chapter 4 and Chapter 6, as well as, the manuscripts in prep cited above.

3.2 Experimental Procedures

3.2.1 Baculovirus Expression and Purification of Full–Length USP7 ($USP7_{FL}$)

Disclaimer: With the exception of the expression from the baculovirus expression system, this section was performed by a previous post–doc in the lab. The data included within this section is the extent of what was provided once I took over the project.

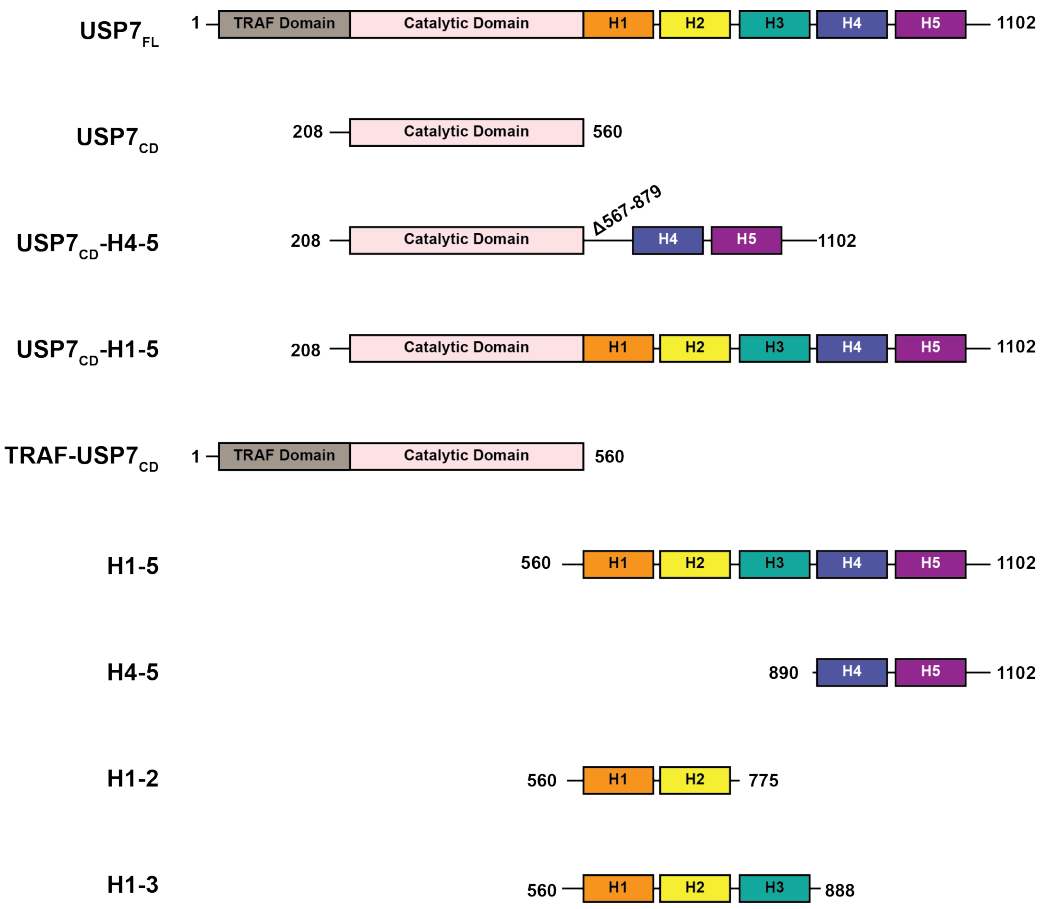


Fig. 3.1. **Truncations of USP7.** Depicted above are the nine constructs that were utilized for the studies described within Chapter 4 and Chapter 6. For each truncation, the residue range is reported as well as the domains encompassed within that range are depicted.

Biotechnology company, BioBasic Inc., codon optimized human $USP7_{FL}$ (1–1102) for baculovirus expression, and performed the cloning of the gene into a modified pVL–1392 vector containing an N–terminal 10His, green fluorescent protein (10x His–GFP) tag. The DNA sequence was verified by DNA sequencing at the Purdue University Genomics Core Facility. A culture volume of 200 mL was used for $USP7_{FL}$ expression by following the BD BaculoGold system in *Sf9* identically to as described in Section 2.3.7. The *Sf9* cells expressing $USP7_{FL}$ were harvested by centrifugation with a Sorvall RC 6+ ThermoScientific Centrifuge with a Fiberlite F9–4x1000y rotor (3000x g for 20 minutes at 4 °C) after 3 days of expression. Approximately 2 g of cell pellet was obtained from the 200 mL of *Sf9* culture under the given expression conditions. The cells were resuspended in 15 mL of Buffer α (20 mM HEPES pH 7.5, 500 mM NaCl, and 10 mM Imidazole) and lysed, on ice, by sonication with a Branson Digital Sonifier at 50% amplitude alternating five seconds on and five seconds off for eight pulses total. The lysate was clarified in Oak Ridge tubes with a Beckman Coulter Avanti J–30I centrifuge in the JA–17 rotor (80,000 x g for one hour at 4 °C).

Nickel Batch Affinity Chromatography

The 15 mL of clarified lysated was bound to 0.5 mL of Ni–NTA resin (GE Healthcare), and unbound protein was washed off with 10 mL of Buffer α. The resin was washed with 10 mL of Buffer β (20 mM HEPES pH 7.5, 500 mM NaCl, 50 mM Imidazole) to remove weakly bound non–specific proteins. $USP7_{FL}$ was eluted with 3 mL of Buffer δ (20 mM HEPES pH 7.5, 500 mM NaCl, and 450 mM Imidazole). The 3.6 mg of eluted $USP7_{FL}$ fusion protein was then digested with PreScission protease to remove the 10x–His–GFP tag at a 1:1 ratio of $USP7_{FL}$ to PreScission Protease. The reaction was allowed to proceed overnight at 4 °C. PreScission protease was prepared as described in Section 3.2.2.

Gel–Filtration Chromatography

As the final step of purification, the cleaved $USP7_{FL}$ protein was purified by a HiPrep 300 16/60 Sephacryl S300 prep grade column equilibrated with Buffer γ (20 mM HEPES pH 7.5, 200 mM NaCl, 10 % glycerol, and 10 mM βME) at a flow rate of 0.5 mL/min with 5 mL fractions collected throughout the run. The pooled sample was flash frozen in liquid nitrogen and stored at -80 °C.

3.2.2 Preparation of PreScission Protease

Expression of PreScission Protease

The expression plasmid pGEX–6P encoding for GST–PreScission protease, as well as the expression and purification protocol, was gifted to us by Professor Jue Chen at Purdue University (now at Rockefeller). The plasmid was transformed by electroporation into *E. coli* BL21 (DE3) cells, and streaked onto LB plates supplemented with 50 μg/mL Carbenicillin. From the colonies that grew, a slurry was made by adding 5 mL of LB media to the plate and gently swirled. The slurry was then added to 100 mL of LB media supplemented with 50 μg/mL Carbenicillin. The culture was grown at 37 °C with shaking at 200 rpm in an Infors Multitron incubated shaker until the cell density resembles that of a normal culture grown overnight (A_{600} ∼6). A 20 mL aliquot of the culture was used to inoculate 1 L of LB media supplemented with 50 μg/mL of Carbenicillin in a Fernbach flask. The culture was grown at 37 °C until A_{600} ∼0.6, about one to two hours. The culture was then cooled to 30 °C, and GST–PreScission Protease expression was induced with IPTG at a final concentration of 0.5 mM from a 1 M stock. The culture was incubated with shaking at 30 °C for 15–17 hours for expression. The cells were harvested with a Sorvall RC 6+ ThermoScientific Centrifuge with a Fiberlite F9–4x1000y rotor at 3000 x g for 15 minutes at 4 °C, and the cell pellet was placed at -80 °C for storage.

Purification of PreScission Protease

The cell pellet containing expressed GST–PreScission was removed from the -80 °C and thawed on ice. The cells were resuspended in 20 mL of PreScission Lysis Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM βME, 5 mM EDTA, 1 mM PMSF, lysozyme, and 3 μg/mL DNAse) and lysed by sonication with a Branson Digital Sonifer at 60% amplitude for three minutes in a metal beaker surrounded by an ice/ethanol bath. Lysate was centrifuged with a Beckman Coulter Avanti J–30I centrifuge in a JA–17 rotor at 80,000 x g for one hour at 4 °C. The supernatant was separated from the insoluble portion, and mixed with 5 mL of Glutathione Sepharose resin equilibrated with Wash Buffer (20 mM HEPES pH 7.5, 5 mM βME, and 5 mM EDTA) in a 50 mL Falcon tube. The mixture was incubated at 4 °C with gentle rocking for two hours. The GST–PreScission protease bound Glutathione Sepharose resin was centrifuged at 500 x g for five minutes and the supernatant was decanted. The resin was washed with a total of 300 mL of wash buffer in 50 mL aliquots. Each aliquot was added to the resin and and incubated with gentle rocking at 4 °C for 10 minutes, then centrifuged with a Sorvall RC6+ ThermoScientific Centrifuge with a Fiberlite F13–14 x 50cy fixed angle rotor at 500 x g for five minutes. This was repeated until all 300 mL of wash buffer had been utilized. To elute GST–PreScission Protease from the Glutathione Sepharose resin 50 mL of Elution Buffer (50 mM HEPES pH 8, 150 mM NaCl, 5 mM βME, and 10 mM reduced Glutahione) was added to the resin, and incubated for 30 minutes with gentle rocking at 4 °C. The sample was again centrifuged with the Sorvall RC6+ ThermoScientific Centrifuge with a Fiberlite F13–14 x 50cy fixed angle rotor at 500 x g for five minutes. The supernatant containing the eluted PreScission protease was collected, and dialyzed against two liters of dialysis buffer (50 mM HEPES pH 8, 150 mM NaCl, 1 mM DTT, and 10 mM EDTA) at 4 °C overnight using SnakeSkin Dialysis Tubing (10,000 MWCO). The following morning, the protein was analyzed by SDS–PAGE and Microassay Bio–Rad Bradford Protein Assay. To store, 20% glycerol was added directly to the sample, the protein was the

diluted to 1 mg/mL with dialysis buffer, and 1 mL aliquots were made. The aliquots were flash frozen in liquid nitrogen and stored at -80 °C.

3.2.3 *E. coli* Expression and Purification of USP7 Truncations

Truncations of USP7 were expressed and purified from *E. coli* to elucidate the mechanism by which USP7 is internally activated (Chapter 4) as well as evaluate the mode of inhibition by identified USP7 specific inhibitors (Chapter 6). A total of 8 truncations were made: the catalytic domain ($USP7_{CD}$, residues 208–560), catalytic domain fusion with HUBL domains 4 and 5 (residues 209–1102 Δ567–879, $USP7_{CD}$–H4–5), USP7 without the N–terminal TRAF domain (residues 209–1102, $USP7_{CD}$–H1–5), TRAF and catalytic domain of USP7 (residues 1–560, TRAF–$USP7_{CD}$), HUBL domains 1–5 (H1–5, residues 560–1102), HUBL domains 4–5 (H4–5, residues 890–1102), HUBL domains 1–2 (H1–2, residues 560–775), and HUBL domains 1–3 (H1–3, residues 560–888). Each truncation was cloned, expressed and purified using the following general procedure. Nuances that pertained to individual truncations are described within their individual results section.

Expression Construct Design

The desired amino acid residues for the $USP7_{CD}$, $USP7_{CD}$–H4–5, $USP7_{CD}$–H1–5, TRAF–$USP7_{CD}$, H1–5, and H4–5 truncations (illustrated in Figure 3.1) was codon optimized by BioBasic Inc for expression within *E. coli*. Each gene was then synthesized by BioBasic Inc with a N–terminal 6His tag and a Tobacco Etch Virus (TEV) cleavage site. Finally, BioBasic Inc inserted each synthesized gene into the pET–11a vector between *NdeI* and *BamHI* restriction sites.

Truncation Protein Expression

For protein expression, *E. coli* BL21–DE3 electro–competent cells were transformed by electroporation with the pET–11a plasmid containing the truncation gene of interest. The transformed cells were gently resuspended in SOP media, and allowed to outgrow for one hour at 37 °C in a 14 mL round bottom culture tube shaking in an Infors Multitron incubated shaker at 200 rpm. An aliquot of 100 μL of the cells were then spread onto LB agar supplemented with 50 μg/mL of carbenicillin. A single colony was picked from the agar plate to innoculate 150 mL of LB media supplemented with 50 μg/mL carbenicillin. The cells were grown overnight at 37 °C. Three one–liter cultures of LB media supplemented with 50 μg/mL carbenicillin were prepared. Each flask was innoculated with 25 mL each of the overnight culture. The cultures were grown to an A_{600} of $\sim$0.6. Cells were cooled to 18 °C, and expression was induced with a final concentration 0.25 mM IPTG from a 1 M stock. The cultures were incubated with shaking in the Infors Multitron incubated shaker at 18 °C for 18 hours. The cells were harvested by centrifugation with a Sorvall RC6+ Thermo-Scientific Centrifuge with a Fiberlite F9–4x1000y rotor (3000 x g for 20 minutes at 4 °C). The cell pellet weight was noted, and 5 mL of Buffer A (10 mM Imidazole, 10 mM Tris pH 7.5, 5 mM βME, 500 mM NaCl, and 5% glycerol) for every one gram of cell pellet was used to resuspend the cells. The resuspended cells were supplemented with 50 μg/mL DNase and 200 μg/mL lysozyme. The cells were lysed with a Branson Digital Sonifier at 60% amplitude pulsing for five seconds on and nine seconds off for a total of 12 minutes on. The lysate was clarified with the same centrifuge with the Fiberlite F13–14 x 50cy fixed angle rotor (28,880 x g for 20 minutes at 4 °C), and the supernatant was filtered through a 0.45 μm filter.

Nickel Affinity Chromatography and TEV Cleavage

The clarified lysate was loaded onto a 5 mL HisTrap column at 2 mL/min (GE Healthcare) equilibrated with Buffer A. Unbound protein was washed off with 50

mL of Buffer A at a flowrate of 2.5 ml/min. A 100 mL linear gradient of 0–100% of Buffer B (450 mM Imidazole, 10 mM Tris pH 7.5, 5 mM βME, 500 mM NaCl, and 5% glycerol) was run to elute bound proteins. Fractions (5 mL) were collected throughout the gradient. Fractions were pooled based on concentration, SDS–PAGE analysis, and specific activity measurements (if the truncation contained the catalytic domain). The pooled protein concentration was measured by Bradford assay and the protein was combined at a 1:25 ratio of TEV protease to truncation and dialyzed in Buffer A without imidizole overnight at 4 °C to remove the 6His tag.

Reverse Nickel Affinity Chromatography

The TEV protease digested sample containing the cleaved 6His tag was loaded onto the 5 mL HisTrap column previously equilibrated with Buffer A, at a flowrate at 2 mL/min. The flowthrough was collected either as 5 mL fractions or as a pool. Unbound sample was washed with 40 mL of Buffer A and a linear gradient of 0–100% of Buffer B was run over 100 mL at a flow rate at 2.5 mL/min. The sample was pooled based on concentration, SDS–PAGE analysis, and specific activity measurements (when appropriate).

Gel–Filtration Chromatography

The pooled sample from the Reverse Nickel Affinity Chromatography was concentrated using ultrafiltration (Millipore) with an appropriate molecular weight cut off membrane to the truncation being purified to < 2 mL. The concentrated samples was injected onto 300 mL GE Healthcare size–exclusion chromatography column previously equilibrated with Buffer C (10 mM Tris pH 7.5, 100 mM NaCl, 5% Glycerol, and 5 mM βME). The type of size–exclusion chromatography column utilized was determined based on the truncations size: Superdex 75 ($USP7_{CD}$, $USP7_{CD}$–H4–5, TRAF–$USP7_{CD}$, H1–5, H4–5, H1–2, and H1–3) or Superdex 200 ($USP7_{CD}$–H1–5). All columns were run at a flowrate of 1 mL/min, and 5 mL fractions were collected

throughout. Fractions to be pooled were determined by concentration, SDS-PAGE analysis, and specific activity measurements (if applicable). The pooled fractions were flash frozen in liquid nitrogen, and stored at -80 °C.

3.2.4 Expression and Purification of USP7 sans the TRAF domain ($USP7_{CD}$–H1–5)

The expression construct design and protein expression of $USP7_{CD}$–H1–5 was performed identically to as described in Section 3.2.3.

Nickel Affinity Chromatography

The clarified lysate of $USP7_{CD}$–H1–5 was loaded onto a 5 mL HisTrap column at 2 mL/min (GE Healthcare) equilibrated with Buffer A. Unbound protein was washed off with 50 mL of Buffer A at a flowrate of 2.5 mL/min. A 100 mL linear gradient of 0–100% of Buffer B was run to elute bound proteins. Fractions (5 mL) were collected throughout the gradient. Fractions were pooled based on concentration, SDS–PAGE analysis, and specific activity measurements.

Superdex 200 Chromatography and TEV Cleavage

The pooled sample was concentrated using ultrafiltration with a 50,000 MWCO membrane (Millipore) to < 3 mL, and injected onto a 300 mL Superdex 200 (GE Healthcare) size–exclusion chromatography column previously equilibrated with Buffer C. The column was run at 1 mL/min, and 5 mL fractions were collected throughout the run. To determine which fractions to pool, peak fractions were analyzed by Bradford assay, SDS–PAGE, and specific activity measurements. The pool was combined in a 1:8 ratio of TEV:$USP7_{CD}$–H1–5 and stored overnight at 4 °C for removal of the 6His tag.

MonoQ Chromatography

The TEV protease digested $USP7_{CD}$–H1–5 sample was diluted in Buffer D (10 mM Tris pH 7.5, 5 % glycerol, and 5 mM βME) 1:3 to reduce the sodium chloride concentration to $<$ 20 mM. The diluted sample was concentrated to $<$ 50 mL with ultrafiltration with a 50,000 MWCO membrane (Millipore). The concentrated sample loaded onto an 8 mL MonoQ column (GE Healthcare) equilibrated in Buffer D at a flowrate of 1 mL/min. Unbound sample was washed with 75 mL of Buffer D at 1 mL/min. A linear gradient of 0–100% of Buffer E (10 mM Tris pH 7.5, 500 mM NaCl, 5 % glycerol, and 5 mM βME) was run at 2.5 mL/min over 100 mL, and 5 mL fractions were collected throughout the run. Fractions were analyzed by concentration, SDS–PAGE, and specific activity measurements, those containing $USP7_{CD}$–H1–5 of the highest purity and activity were pooled, flash frozen in liquid nitrogen, and stored at -80 °C.

3.2.5 Sub–Cloning and Construct Design for H1–2

Primers were designed to amplify the bases encoding for the USP7 HUBL domains 1 and 2 (residues 560–775) from the H1–5–pET–11a vector for insertion into the pEV–L8 vector by ligation independent cloning techniques. The forward and reverse primers used are listed below from 5' to 3'. The bases in bold are bases complimentary to the pEV–L8 vector that allows for homologus recombination. Forward: **GAGAACCTGTACTTCCAATCCAAT**GAAGCACATCTGTACATGCAG. Reverse: **TCGAATTCGGATCCGTTATCCACTTCCAAT**CGTTTTCCGGGTC GTC. The master mix for a single 50 μL PCR reaction was as described in Table 3.1. The PCR reaction was performed with a Bio–Rad MJ Mini Personal Thermal Cycler with the reaction scheme described in Table 3.2.

The PCR products were analyzed with a 1% agarose gel supplemented with 1x GelRed dye that was run for 40 minutes at 100 volts. The successful reactions were pooled and the cleaned using the Qiagen PCR clean–up protocol. The PCR ampli-

Table 3.1
PCR composition for the sub–cloning of the *H1–2* gene

Component	Final Concentration	Volume
5x Phusion HF Buffer	1x	5 μL
10 mM dNTPs	0.2 mM	0.5 μL
10 μM Forward Primer	0.25 μM	5 μL
10 μM Reverse Primer	0.25 μM	5 μL
USP17–pET11a plasmid	50–100 ng	–
Phusion High–Fidelity DNA Polymerase	0.5 units	0.4 μL
Sigma–Water	N/A	to 50 μL

Table 3.2
PCR composition for the sub–cloning of the *H1–2* gene

Step	Temperature	Time
Initial Denaturation	95 °C	3 minutes
Denaturation	95 °C	30 seconds
Annealing	48 °C	30 seconds
Extension	72 °C	1.5 minutes
Repeat steps 2-4 for 30 cycles		
Completion	72 °C	5 minutes
Hold	4 °C	

con was then inserted into the pEV–L8 vector by ligation–independent recombinant cloning using linearized pEV–L8 vector, digested by restriction enzyme SspI, and XL1-Blue supercompetent cells, as described in Chapter 2. The construct was verified by DNA sequencing at the Purdue University Genomics Core Facility.

3.2.6 Sub–Cloning and Construct Design for H1–3

Primers were designed to amplify the bases encoding for the USP7 HUBL domains 1 and 3 (residues 560–888) from the pET–11a vector encoding for H1–5. This region was inserted into the pEV–L8 vector by ligation independent cloning techniques, as described in Section 3.2.5. The forward and reverse primers used are listed below from 5' to 3'. The bases in bold are bases complimentary to the pEV–L8 vector that allow for homologus recombination. Forward: GAGAACCTGTACTTCCAATCCAATGAAGCACATCTGTACATGCAG. Reverse: TCGAATTCGGATCCGTTATCCACTTCCAATGAAGTCCGTGATTTTCATTTTCAG. The PCR master mix, PCR reaction, and the homologus recombination cloning was identical to what was described in Section 3.2.5. The construct was verified by DNA sequencing at the Purdue University Genomics Core Facility.

3.2.7 Expression and Purification of Tobacco Etch Virus (TEV) Protease

Expression of TEV Protease

The expression plasmid (pRK793) encodes for the Maltose Binding Protein (MBP)–6His tag– TEV(S219V)– polyarginine tag TEV construct. The S192V mutation eliminates the internal autoinhibition regulation of TEV protease. Between the MBP and N–terminal histidine tag is a TEV protease recognition site (ENLYFQG/S). During expression, soluble TEV protease will cleave itself from the MBP fusion protein exposing the N-terminal 6His tag for use during purification. The plasmid is transformed into *E. coli* BL21–RIL (codon +) cells, and streaked onto LB plates supplemented with 50 μg/mL Carbenicillin and 30 μg/mL Chloramphenicol. The plates were allowed to grow overnight at 37 °C. A single colony was picked and used to innoculate 30 mL of LB supplemented with 50 μg/mL Carbenicillin and 30 μg/mL Chloramphenicol. The started culture was grown at 37 °C shaking at 200 rpm overnight. Three Fernbach flasks containing 1 L of LB media supplemented with 50 mg/mL

Carbenicillin and 30 mg/mL Chloramphenicol were prepared. Each flask was innoculated with 10 mL of the starter culture, and grown at 37 °C to an A_{600} of 0.5. The flasks were then induced with 1 mM IPTG, and allowed to express for four hours at 30 °C. Cells were harvested by centrifugation with the Sorvall RC6+ ThermoScientific Centrifuge with the Fiberlite F9–4x1000y rotor at 4000 x g for 20 minutes at 4 °C. The pelleted cell weight was 12.81 g for 3 L of culture. The pelleted cells were resuspended in 25 mL of TEV Lysis Buffer (20 mM Tris pH 8.0, 100 mM NaCl, 10 % glycerol, and 10 mM βME) and lysed by sonication with a Branson Digital Sonifer at 55% amplitude pulsating for 10 seconds on and 10 seconds off for a total of seven minutes. During sonication, the cells were in a metal beaker surrounded by an ice/ethanol slurry to keep sample from overheating. The lysed cells were centrifuged in the same centrifuge with the Fiberlite F12–14 x 50cy fixed angle rotor at 28,880 x g for 25 minutes at 4 °C.

Purification of TEV Protease

The clarified lysate containing active TEV protease cleaved from MBP was then loaded onto a 5 mL HisTRAP column at 2 mL/min (GE Healthcare) equilibrated with TEV Buffer A (20 mM Tris pH 8.0, 100 mM NaCl, 10% Glycerol, 10 mM βME, and 10 mM Imidizole). The column was washed with 10 column volumes (50 mL) of TEV Buffer A to remove nonspecific proteins. TEV protease was eluted from the column with a gradient from 0–100% TEV Buffer B (20 mM Tris pH 8.0, 100 mM NaCl, 10% Glycerol, 10 mM βME, and 250 mM Imidizole) over 20 column volumes (100 mL). During the gradient 5 mL fractions were collected. The fractions were analyzed by 12.5% SDS–PAGE gel, and pure fractions were pooled. The pooled TEV sample was immediately dialyzed for three to four hours against 2 L of TEV Dialysis Buffer (20 mM Tris pH 8.0, 100 mM NaCl, 10% Glycerol, 10 mM βME, and 1 mM EDTA) to remove the imidizole and add EDTA to the buffer. The dialysis buffer was refreshed with 2 L of fresh TEV Dialysis Buffer and the TEV protease sample

was dialyzed overnight. The following morning the dialysis buffer was refreshed again with 2 L of fresh TEV Dialysis Buffer, and the TEV protease sample was dialyzed for an additional four hours. The TEV protease sample was then concentrated using ultrafiltration with a 3,000 MWCO membrane (Millipore) to a concentration of 1 mg/mL. TEV was then flash frozen in 1 mL aliquots in liquid nitrogen and stored at -80 °C. On average, one can expect 50 mg of TEV protease per liter of cell culture. The TEV purification is summarized in Figure 3.2.

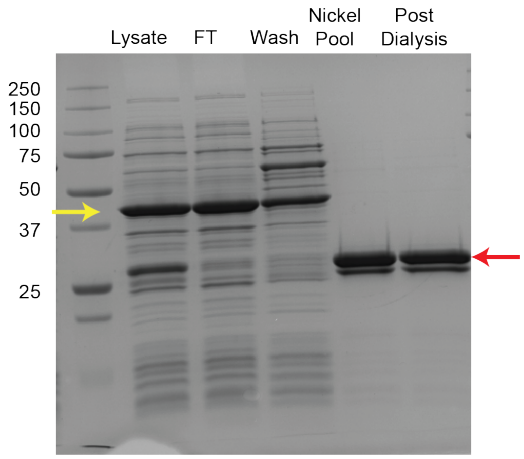


Fig. 3.2. **Summary of the TEV Protease Purification.** 12.5% SDS–PAGE summary of the TEV protease purification. The red arrow indicates the purified TEV protease at the expected size of 27 kDa. The yellow arrow is the cleaved Maltose Binding Protein. Molecular weight marker sizes as indicated.

3.2.8 Kinetic Assays of USP7

The specific activity was measured throughout each purification of the USP7 constructs containing the catalytic domain ($USP7_{FL}$, $USP7_{CD}$, $USP7_{CD}$–H4–5, $USP7_{CD}$–H1–5, and TRAF–$USP7_{CD}$) using the fluorogenic synthetic substrate ubiquitin–rhodamine 110 (Ub–Rho110, Boston Biochem). Ub–Rho110 is internally quenched when the rhodamine 110 molecule is conjugated to the ubiquitin moiety by a peptide bond. Upon cleavage of the peptide bond the quenching is released, and fluorescence from the

rhodamine 110 can be measured. This substrate works as a generic substrate for many USP enzymes (Figure 2.5). The Ub–Rho110 is provided from Boston Biochem at a stock concentration of 250 μM in 100% DMSO which is normalized to 1 μM in Substrate Buffer (50 mM Tris pH 7.5, 1 mM EDTA, 100 mM NaCl, and 0.05% CHAPS), and diluted into each assay at a final concentration of 500 nM. The protein concentration was measured by the Microassay Bio–Rad Bradford Protein Assay and was normalized in Assay Buffer (50 mM Tris pH 7.5, 0.1 mg/mL BSA, and 0.01% Trition X–100) to a working concentration of 250 nM for TRAF–USP7$_{\mathrm{CD}}$, 10 nM for USP7$_{\mathrm{CD}}$ and USP7$_{\mathrm{CD}}$–H4–5, or 2 nM for USP7$_{\mathrm{CD}}$–H1–5. First, 15 μL of the working concentration of Ub–Rho110 (1 μM) was added to a Costar 96 half–volume black plate, and the background fluorescence was read using a BioTeK Synergy H1 plate reader every 10 seconds at an excitation λ of 485 nm and emission λ of 528 nm for two minutes. Following the background read, 15 μL of protein was added to the plate (a 2–fold dilution), the plate was mixed for 10 seconds, then the assay was read under the same conditions for 30 minutes total or until the reactions started to plateau.

Determination of the Extinction Coefficient of the Substrate

The extinction coefficient of Ub–Rho110 was determined anew each day of each purification to account for freeze–thaw effects. To determine the extinction coefficient, five serial dilutions of Ub–Rho110 were made to the working concentrations of 0.03 μM to 1 μM. First, 15 μL of each dilution of Ub–Rho110 was added to the plate. Second, 15 μL of a highly concentrated USP stock was added the plate. The plate was mixed for 10 seconds and the fluoresence was read with the same program described for the specific activity assays. The reactions were allowed to plateau, and the maximum artificial fluorescence units (AFU) was measured for each concentration. The measured AFU's were plotted against the corresponding Ub–Rho110 concentration to build a standard curve. The resulting plot was fit using linear regression,

and the resulting slope corresponds to the extinction coefficient in units of AFUs per concentration of Ub–Rho110 (AFU/μM).

Specific Activity Analysis

The initial slope of each specific activity reaction in units of fluorescence per unit time (AFU/min) was converted to the amount of product produced per unit time (μM/min) with the extinction coefficient of the substrate. Units, as defined by μmol/min^{-1}, was obtained by dividing by the total assay volume. The specific activity of the sample, Units/(mg of enzyme), was determined by dividing the Units by the milligrams of protein present in the assay. The fractional recovery of activity was determined by dividing the total units of the sample by the total units of the lysate and multiplying by 100 to obtain a percentage. The fold in purification after each chromatography step was determined by dividing the specific activity of each sample by the specific activity of the lysate.

3.2.9 Measuring Protein Concentration

Microassay Bio–Rad Bradford Protein Assay

The Microassay Bio–Rad Bradford Protein Assay was used throughout each purification to estimate the concentraiton of the total protein in the sample. The coomassie dye was prepared by diluting the Bio–Rad 5x dye concentrate to 1x with nanopure water. A standard curve was prepared using Bio–Rad lyophilized bovine serium albumin (BSA) that had been resuspended to a stock concentration of 1.36 mg/mL in nanopure water. This stock was then diluted to a working stock of 0.5 mg/mL in the same purification buffer the protein sample to be measured was currently residing. A standard curve of BSA was measured at 0, 0.125, 0.25, 0.375, and 0.5, mg/mL in a 96–well clear, flat–bottom microtiter plate. The protein samples were added to the plate in a volume of 1, 5, or 10 μL depending on the concentration of the sample to

ensure the staining fell within the BSA standard curve. If 1 μL or 5 μL of sample was used the dilution factor of 10 or 2, respectively, was taken into account for the final calculation. To each well 200 μL of the 1x Bio–Rad coomassie dye was added to each sample. Using a microplate mixer, the plate was mixed thoroughly for 30 seconds and any bubbles were popped. Using the BioTek H1 Synergy plate reader, the absorbance of each well was read at a λ of 595 nm. The absorbance of the BSA standard curve was plotted in terms of the corresponding concentration, and fit using linear regression from that the protein concentration of each sample was estimated.

Determination of the Extinction Coefficient of USP7 Truncations

The extinction coefficient was determined for each purified truncation as described in Section 2.3.9, and are summarized in Table 3.7. Samples labeled with an asterisk indicate that the purity was too poor or the yield was too low for proper determination of the extinction coefficient by this method. For those cases the values displayed were determined computationally by the bioinformatics resource, ExPASy [52].

3.2.10 SDS–PAGE Analysis

Protein sample concentration was measured by a Microassay Bio–Rad Bradford Protein Assay. SDS–PAGE samples were prepared by mixing 5–10 μg of protein sample with 4 μL of 5x SDS Sample Loading Buffer (0.2 M Tris–HCl pH 6.8, 10% SDS, 10% glycerol, 0.02% Bromophenol blue, and 5 mM βME). Samples were heated at 95 °C for two minutes immediately prior to loading into the SDS–PAGE gel. The gel was run at 200 volts for 40 minutes. Each USP7 construct was analyzed by 10% or 12.5% SDS–PAGE gel depending on the size.

3.3 Results

3.3.1 Expression and Purification of $USP7_{FL}$

Figure 3.3 displays the SDS–PAGE analysis of purified $USP7_{FL}$. From this preparation, approximately 1 mg of $USP7_{FL}$ was obtained from 200 mL of *Sf9* culture. The specific activity of $USP7_{FL}$ was only recorded in the notes provided to me for the final enzyme pool. The specific activity was determined to be 0.1 Units/mg (a unit is defined as $\mu mol/min^{-1}$).

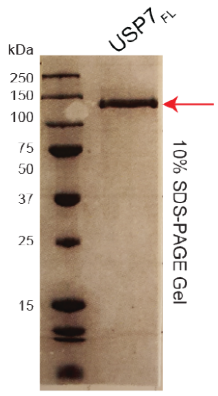


Fig. 3.3. **SDS–PAGE analysis of purified $USP7_{FL}$.** 10% SDS–PAGE analysis of the final pool of the purified $USP7_{FL}$. Red arrow indicates expected size of $USP7_{FL}$, 125 kDa. Molecular weight marker sizes as indicated.

3.4 Expression and Purification of USP7 Catalytic Domain ($USP7_{CD}$)

The expression and purification of $USP7_{CD}$ was performed identically to as described in Section 3.2.3. Approximately 10 g of harvested cells were obtained from the 3 L culture. Figure 3.4 displays the chromatograms and SDS–PAGE analysis associated with the purification of $USP7_{CD}$. Figure 3.4A displays the nickel affinity chromatography where fractions 5–10 were pooled and digested with TEV protease to remove the 6His tag. Figure 3.4B represents the reverse nickel affinity chromatog-

raphy where $USP7_{CD}$ was washed out from the column with the unbound proteins. Figure 3.4C displays the final Superdex 75 chromatography step where fractions 31–34 were pooled. Approximately 67 mg of pure protein was obtained from the 3 L expression. A summary of the percent enzyme yield, total activity, and fold–purification after each chromatography step is summarized in Table 3.3.

Table 3.3
Purification summary of $USP7_{CD}$ from 3 L culture of *E.coli* BL21-DE3

Sample	Protein (mg)	Total Activity Units	Specific Activity (Units/mg)	Fold Purification	% Yield
Lysate	474	0.2	0.0005	1	100
Ni–Pool	177	0.02	0.0001	3	104
Reverse Ni–Pool	81	0.1	0.002	4	50
Superdex 75	71	0.2	0.003	6	22

3.4.1 Expression and Purification of $USP7_{CD}$ H4–5 fusion ($USP7_{CD}$–H4–5)

Variations to the expression and purification described in Section 3.2.3 for the preparation of $USP7_{CD}$–H4–5 are as follows. Only 1 L of $USP7_{CD}$–H4–5 was expressed, instead of the described 3 L. During the nickel affinity chromatography, unbound proteins were washed off with only 25 mL of Buffer A instead of the described 50 mL. The pooled fractions from the nickel affinity chromatography was digested with TEV protease at a 1:5 ratio of $USP7_{CD}$–H4–5:TEV instead of the 1:25 ratio to remove the 6His tag.

Approximately 4 g of harvested cells were obtained from the 1 L culture. Figure 3.5 displays the chromatograms and SDS–PAGE analysis associated with the purification of $USP7_{CD}$–H4–5. Figure 3.5A displays the nickel affinity chromatography where

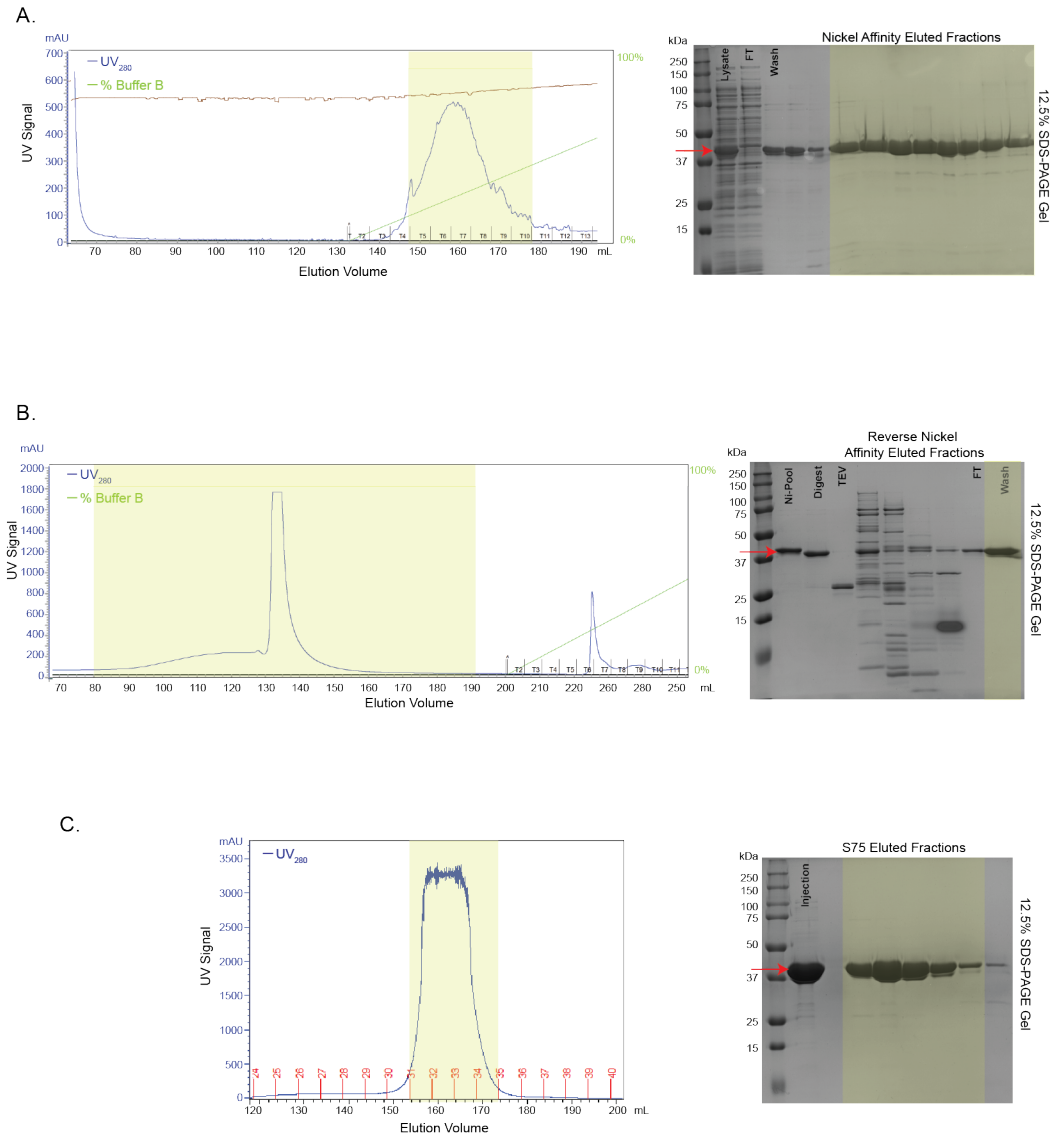


Fig. 3.4. **Purification of USP7$_{CD}$. A.** *Left*–Elution profile from HisTrap (GE Healthcare) column charged with Nickel Sulfate. *Right*–12.5% SDS–PAGE analysis of the eluted fractions, of which the wash and fractions 5–10 were pooled. Red arrow indicates expected size of USP7$_{CD}$. Molecular weight marker sizes as indicated. **B.** *Left*–Elution profile from a 5 mL Hi-Trap (GE Healthcare) column charged with Nickel Sulfate after TEV digest of 6His tag. *Right*–12.5% SDS–PAGE analysis of eluted fractions, of which the wash contained USP7$_{CD}$. **C.** *Left*–Elution profile from a 300 mL Superdex 75 column. *Right*–12.5% SDS–PAGE analysis of eluted fractions, of which 31–34 were pooled.

$USP7_{CD}$–H4–5 eluted at 30% Buffer A, and fractions 13–15 were pooled. Figure 3.5B represents the reverse nickel affinity chromatography where $USP7_{CD}$–H4–5 weakly associated to the column and eluted after only 10% of Buffer A in fractions 9–10. Figure 3.5C displays the final Superdex 75 chromatography step where fractions 24–26 were pooled. Approximately 1.5 mg of pure protein was obtained from the 1 L expression. A summary of the percent enzyme yield, total activity, and fold–purification after each chromatography step is summarized in Table 3.4.

Table 3.4
Purification summary of $USP7_{CD}$–H4–5 from 1 L culture of *E.coli* BL21-DE3

Sample	Protein (mg)	Total activity Units	Specific activity (Units/mg)	Fold Purification	% Yield
Lysate	155	4.6	0.03	1	100
Ni–Pool	16.3	2.1	0.13	4	46
Reverse Ni–Pool	5.3	0.7	0.14	5	16
Superdex 75	1.5	0.7	0.18	6	14

3.4.2 Expression and Purification of USP7 sans the TRAF domain ($USP7_{CD}$–H1–5)

The alternative purification process utilized for the purification of $USP7_{CD}$–H1–5 is described in Section 3.2.4. Approximately 10 g of harvested cells were obtained from the 3 L culture. Figure 3.6 displays the chromatograms and SDS–PAGE analysis associated with the purification of $USP7_{CD}$–H1–5. Figure 3.6A displays the nickel affinity chromatography where fractions 4–10 were pooled, and digested with TEV protease to remove the 6His tag. Figure 3.6B summarizes the Superdex 200 size–exclusion chromatography where fractions 34–36 were pooled. Figure 3.6C displays the MonoQ chromatography where fractions 20–24 were pooled. Approximately 12

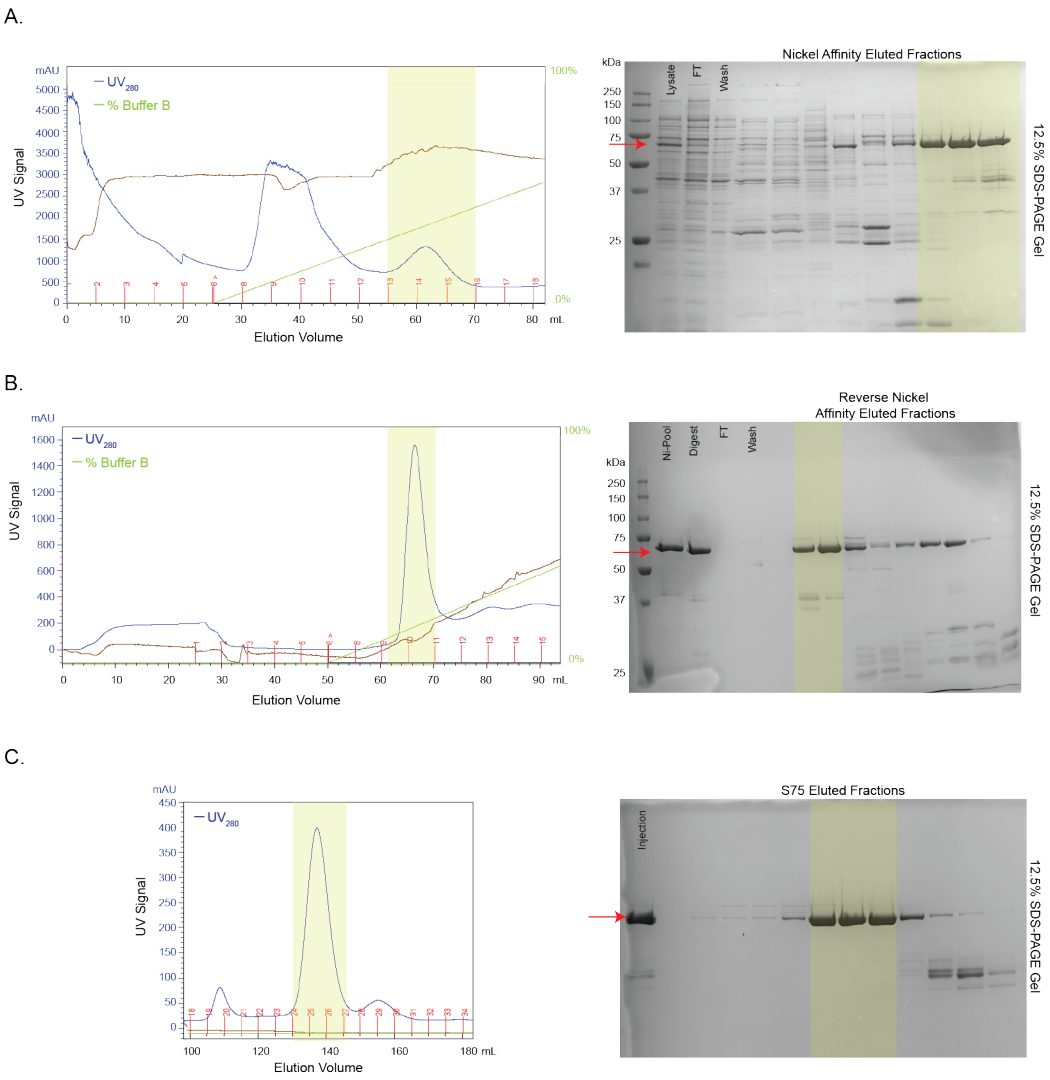


Fig. 3.5. **Purification of USP7$_{CD}$ – H4 – 5. A.** *Left*–Elution profile from a 5 mL Hi-Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*–10% SDS–PAGE analysis of the eluted fractions, of which fractions 13–15 were pooled. Red arrow indicates expected size of USP7$_{CD}$–H4–5, 68 kDa. Molecular weight marker sizes as indicated. **B.** *Left*–Elution profile from a 5 mL Hi-Trap (GE Healthcare) column charged with Nickel Sulfate after TEV digest of 6His tag. *Right*–SDS–PAGE analysis of eluted fractions, of which 9 and 10 contained USP7$_{CD}$–H4–5. **C.** *Left*–Elution profile from a 300 mL Superdex 75 column. *Right*– 10% SDS–PAGE analysis of eluted fractions, of which fractions 24–26 were pooled.

mg of $USP7_{CD}$–H1–5 was obtained from the 3 L culture. A summary of the percent enzyme yield, total activity, and fold–purification after each chromatography step is summarized in Table 3.5

Table 3.5
Purification summary of $USP7_{CD}$–H1–5 from 1 L culture of *E.coli* BL21-DE3

Sample	Protein (mg)	Total Activity Units	Specific Activity (Units/mg)	Fold Purification	% Yield
Lysate	447	16	0.04	1	100
Ni–Pool	103	30	0.3	8	180
Superdex 200	16	5	0.3	8	28
Mono Q	12	3	0.2	7	18

3.4.3 Expression and Purification of TRAF–$USP7_{CD}$

No variations of the Section 3.2.3 protocol were made for the expression and purification of TRAF–$USP7_{CD}$. Approximately 10.5 g of harvested cells were obtained from the 3 L culture. Figure 3.7 displays the chromatograms and SDS–PAGE gels associated with the purification of TRAF–$USP7_{CD}$. Figure 3.7A displays the nickel affinity chromatography of TRAF–$USP7_{CD}$ where fractions 5–12 were pooled. Figure 3.7B summarizes the reverse nickel affinity chromatography where TRAF–$USP7_{CD}$ was collected in the wash. Figure 3.7C displays the final Superdex 75 chromatography step where fractions C3–C5 were pooled. Approximately 3.2 mg of pure protein was obtained from the 3 L expression. A summary of the percent enzyme yield, total activity, and fold–purification after each chromatography step is summarized in Table 3.6.

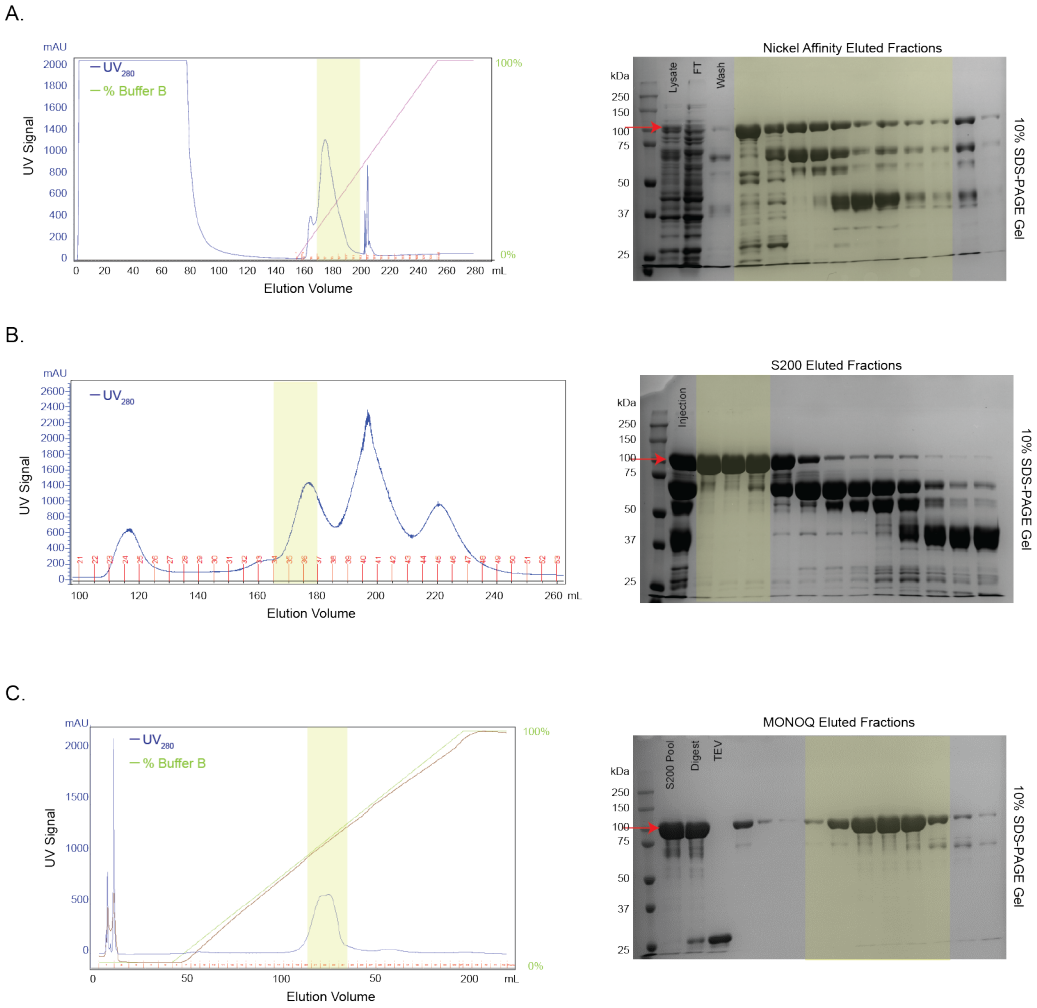


Fig. 3.6. **Purification of USP7$_{CD}$–H1–5. A.** *Left*–Elution profile from 5 mL Hi-Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*–10% SDS–PAGE analysis of the eluted fractions, of which fractions 4–10 were pooled. Molecular weight marker sizes as indicated. Estimated weight of USP7$_{CD}$–H1–5 is 104 kDa. **B.** *Left*–Elution profile from 300 mL Superdex 200 (GE Healthcare) column. *Right*–SDS–PAGE analysis of eluted fractions, of which 35–37 containted USP7$_{CD}$–H1–5. **C.** *Left*–Elution profile from 8 mL MonoQ (GE Healthcare) column. *Right*– 10% SDS–PAGE analysis of eluted fractions, of which fractions 20–24 were pooled.

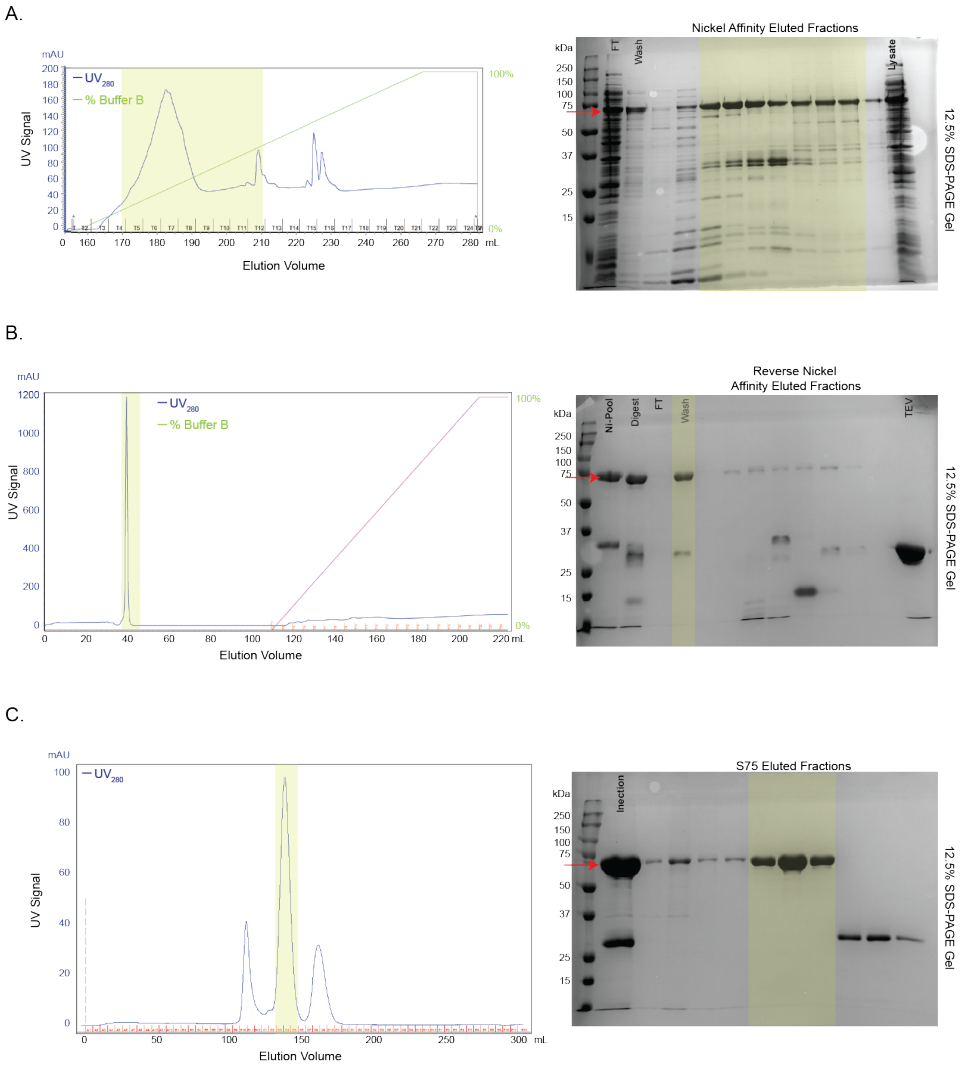


Fig. 3.7. **Purification of TRAF–USP$_{CD}$.** **A.** *Left*–Elution profile from 5 mL Hi–Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*– 12.5% SDS–PAGE analysis of the eluted fractions, of which fractions 5–12 were pooled. Molecular weight marker sizes as indicated. Red arrow indicates estimated molecular weight of TRAF–USP7$_{CD}$, 66 kDa. **B.** *Left*–Elution profile from 5 mL Reverse Hi–Trap (GE Healthcare) column after TEV cleavage of the 6His tag. *Right*–SDS–PAGE analysis of eluted fractions, of which the wash was retained. **C.** *Left*–Elution profile from 300 mL Superdex 75 (GE Healthcare) column. *Right*– 12.5% SDS–PAGE analysis of eluted fractions, of which fractions C3–C5 were pooled.

Table 3.6
Purification summary of TRAF–USP7$_{CD}$ from 3 L culture of *E.coli* BL21-DE3

Sample	Protein (mg)	Total Activity Units	Specific Activity (Units/mg)	Fold Purification	% Yield
Lysate	445	0.3	0.0008	1	100
Ni–Pool	10.7	0.02	0.002	2	6
Reverse Ni–Pool	8.6	0.007	0.003	4	2
Superdex	3.2	0.004	0.002	3	1

3.4.4 Expression and Purification of USP7 HUBL domains 1–5 (H1–5)

Variations to the expression and purification described in Section 3.2.3 for the preparation of H1–5 are as follows. The pooled fractions from the nickel affinity chromatography was digested with TEV protease at a 1:15 ratio of TEV:H1–5 instead of the described 1:25 ratio to remove the 6His tag. During the reverse nickel affinity chromatography step, the column was washed with a total of 120 mL of Buffer A, instead of the described 40 mL.

Approximately 10.5 g of harvested cells were obtained from the 3 L culture. Figure 3.8 displays the chromatograms and SDS–PAGE analysis associated with the purification of H1–5. Figure 3.8A displays the nickel affinity chromatography where fractions 3–16 were pooled, and digested with TEV protease to remove the 6His tag. Figure 3.8B summarizes the reverse nickel affinity chromatography where H1–5 was washed out from the column with the unbound proteins as well as eluted in minimal imidizole in fraction 3. Figure 3.8C displays the final Superdex 75 chromatography step where fractions B10–B11 and were pooled. Approximately 17 mg of pure protein was obtained from the 3 L expression.

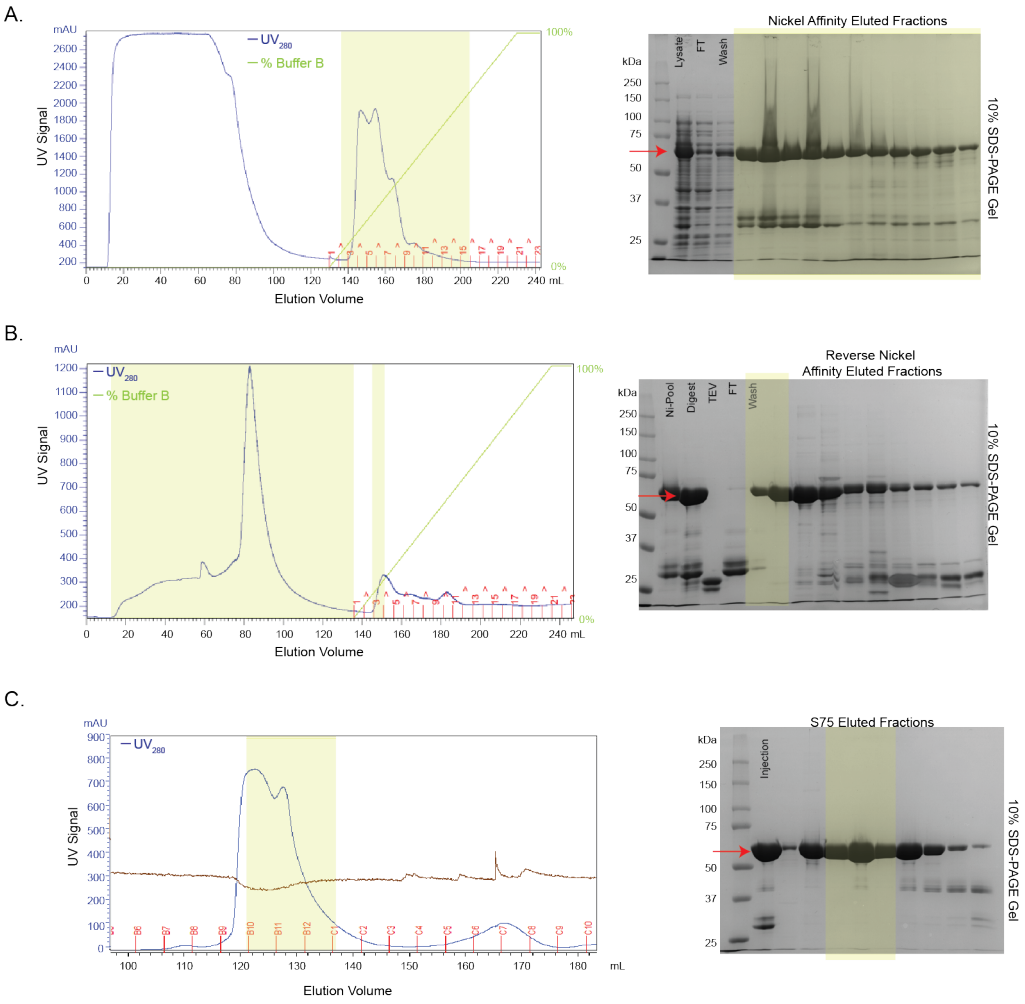


Fig. 3.8. **Purification of H1–5. A.** *Left*–Elution profile from 5 mL Hi–Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*–10% SDS–PAGE analysis of the eluted fractions, of which fractions 3–16 were pooled. Molecular weight marker sizes as indicated. Red arrow indicates estimated molecular weight of H1–5, 64 kDa. **B.** *Left*–Elution profile from 5 mL Reverse Hi–Trap (GE Healthcare) column after TEV cleavage. *Right*–SDS–PAGE analysis of eluted fractions, of which the wash and fraction 3 were pooled. **C.** *Left*–Elution profile from 300 mL Superdex 75 (GE Healthcare) column. *Right*– 10% SDS–PAGE analysis of eluted fractions, of which fractions B10–B11 were pooled.

3.4.5 Expression and Purification of USP7 HUBL domains 4–5 (H4–5)

Variations to the expression and purification described in Section 3.2.3 for the preparation of H4–5 are as follows. The pooled fractions from the nickel affinity chromatography was digested with TEV protease at a 1:10 ratio of TEV:H4–5 instead of the described 1:25 ratio to remove the 6His tag. During the reverse nickel affinity chromatography step, the column was washed with a total of 30 mL of Buffer A, instead of the described 40 mL.

Approximately 14 g of harvested cells were obtained from the 3 L culture. Figure 3.9 displays the chromatograms and SDS–PAGE analysis associated with the purification of H4–5. Figure 3.9A displays the nickel affinity chromatography where fractions A2–B1 were pooled, and digested with TEV protease to remove the 6His tag. Figure 3.9B summarizes the reverse nickel affinity chromatography where H4–5 was washed out from the column with the unbound proteins, as well as eluted with minimal imidizole ($\leq 20\%$) in fractions A1–A9. Figure 3.9C displays the final Superdex 75 chromatography step where fractions C6–C8 and were pooled. Approximately 18 mg of pure protein was obtained from the 3 L expression.

3.4.6 Expression and Purification of USP7 HUBL domains 1–2 (H1–2)

Variations to the expression and purification described in Section 3.2.3 for the preparation of H1–2 are as follows. During the reverse nickel affinity chromatography step, the column was washed with a total of 30 mL of Buffer A, instead of the described 40 mL. The H1–2 purification did not proceed past the reverse nickel chromatography step due to the high level of purity obtained for the H1–2 sample.

Approximately 12 g of harvested cells were obtained from the 3 L culture. Figure 3.10 displays the chromatograms and SDS–PAGE analysis associated with the purification of H1–2. Due to a computer glitch, the chromatogram associated with the SDS–PAGE analysis of the nickel affinity chromatography column displayed in Figure 3.10A was lost. Based on analysis of just the SDS–PAGE gel in Figure 3.10A

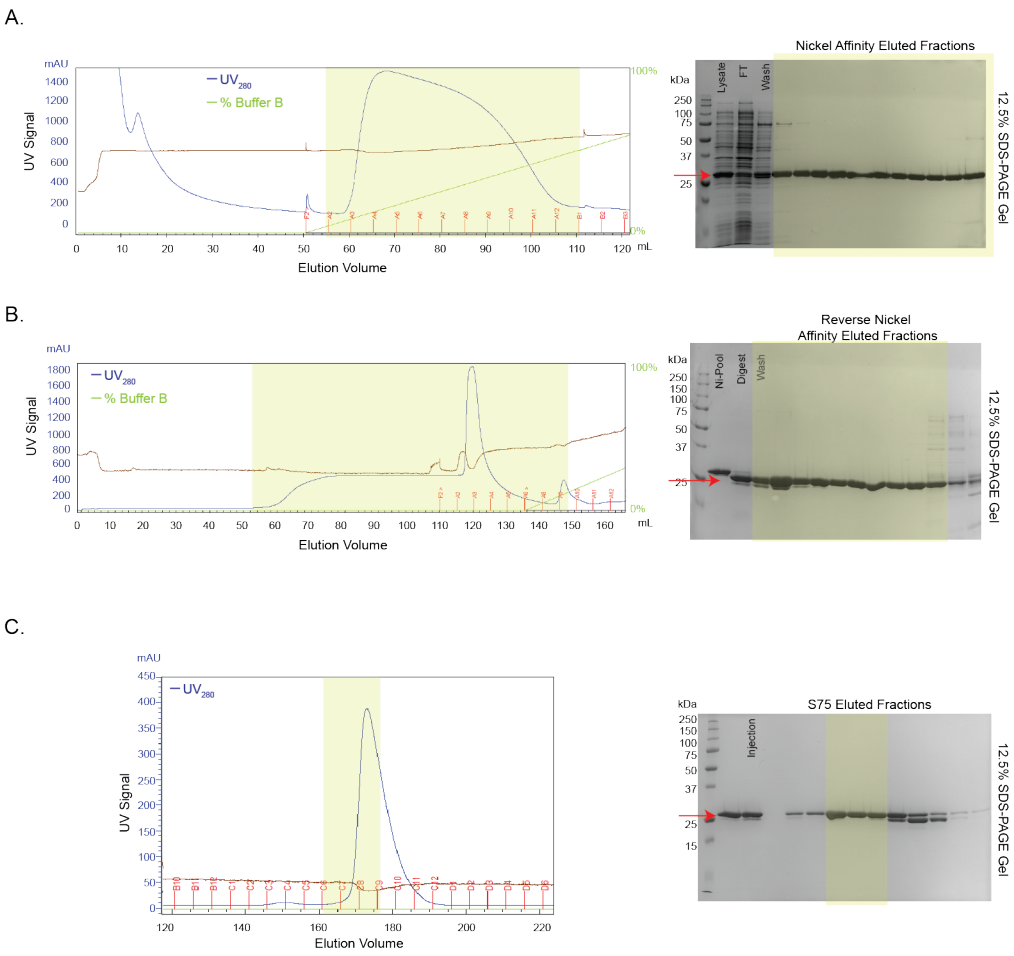


Fig. 3.9. **Purification of H4–5. A.** *Left*–Elution profile from 5 mL Hi–Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*–12.5% SDS–PAGE analysis of the eluted fractions, of which fractions A2–B1 were pooled. Molecular weight marker sizes as indicated. Red arrow indicates estimated molecular weight of H4–5, 25 kDa. **B.** *Left*–Elution profile from 5 mL Reverse Hi–Trap (GE Healthcare) column after TEV cleavage. *Right*–12.5% SDS–PAGE analysis of eluted fractions, of which the flowthrough and fractions A1–A9 were pooled. **C.** *Left*–Elution profile from 300 mL Superdex 75 (GE Healthcare) column. *Right*– 12.5% SDS–PAGE analysis of eluted fractions, of which fractions C6–C8 were pooled.

fractions 3–19 were pooled, and digested with TEV protease to remove the 6His tag. Figure 3.10B summarizes the reverse nickel affinity chromatography where H1–2 was washed out from the column with the unbound proteins as well as eluted with minimal imidizole ($\leq$ 20%) in fractions 5–7. Approximately 132 mg of pure protein was obtained from the 3 L expression.

3.4.7 Expression and Purification of USP7 HUBL domains 1–3 (H1–3)

Variations to the expression and purification described in Section 3.2.3 for the preparation of H1–3 are as follows. For both the nickel affinity and reverse nickel affinity chromatography, the wash step was collected in fractions (5 mL) instead of as a pool. The TEV protease digestion was performed with a 1:15 ratio of TEV:H1–3.

Approximately 12 g of harvested cells were obtained from the 3 L culture. Figure 3.11 displays the chromatograms and SDS–PAGE analysis associated with the purification of H1–3. Figure 3.11A displays the SDS–PAGE analysis of the nickel affinity chromatography where fractions 15–22 were pooled, and digested with TEV protease to remove the 6His tag. Figure 3.11B summarizes the reverse nickel affinity chromatography where H1–3 was present in the flowthrough, as well as washed out in wash fractions 1–3. Figure 3.11C displays the final Superdex 75 chromatography step where fractions C2–C5 were pooled. Approximately 93 mg of pure protein was obtained from the 3 L expression.

3.4.8 Purification Summary

Table 3.7 summarizes the molecular weight, determined extinction coefficient, yield from 1 L of bacterial culture, and the specific activity of the pure protein (if applicable). The highest protein yields were obtained from the inactive HUBL truncations H1–2 and H1–3. The lowest protein yields were of $USP7_{FL}$ from the baculovirus expression system and the TRAF–$USP7_{CD}$ truncation. As reported previously, the specific activity is comparable for $USP7_{FL}$, $USP7_{CD}$–H4–5, and $USP7_{CD}$–H1–5 while

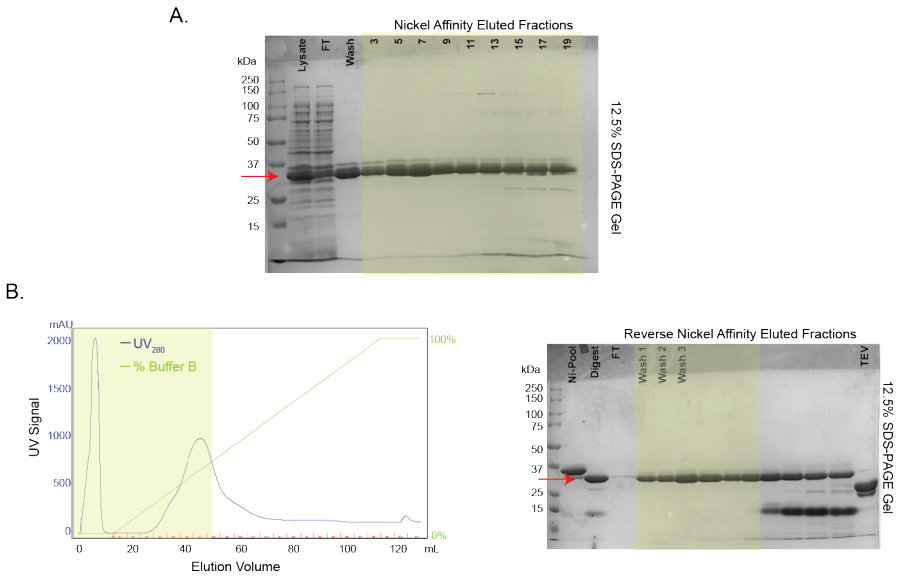


Fig. 3.10. **Purification of H1–2. A.** 12.5% SDS–PAGE analysis of the eluted fractions, of which fractions 3–19 were pooled. Molecular weight marker sizes as indicated. Red arrow indicates estimated molecular weight of H1–2, 25 kDa. **B.** *Left*–Elution profile from 5 mL Reverse Hi–Trap (GE Healthcare) column after TEV cleavage. *Right*–12.5% SDS–PAGE analysis of eluted fractions, of which the wash and fractions 5–7 were pooled.

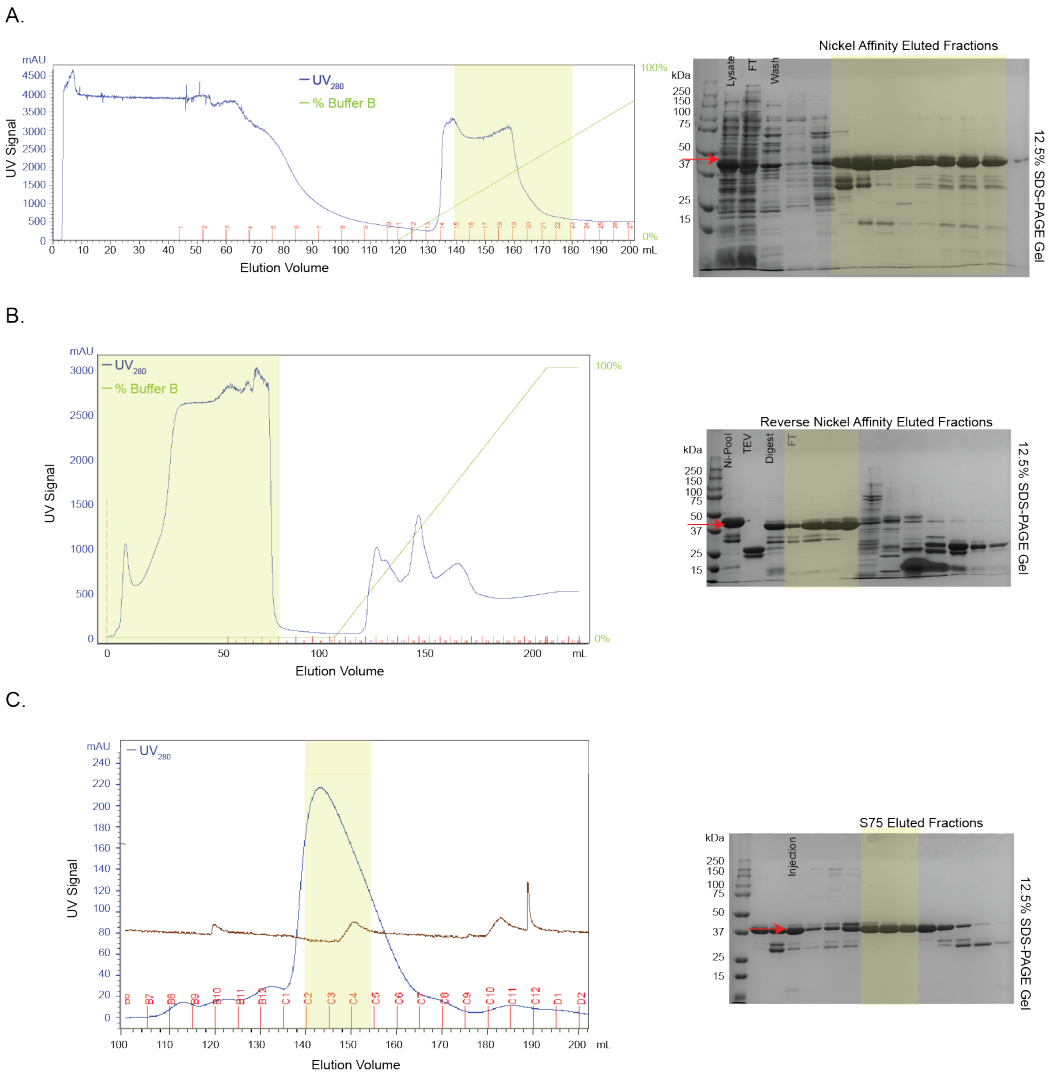


Fig. 3.11. **Purification of H1–3. A.** *Left*–Elution profile from 5 mL Hi–Trap (GE Healthcare) column charged with Nickel Sulfate. *Right*–12.5% SDS–PAGE analysis of the eluted fractions, of which fractions 15–22 were pooled. Molecular weight marker sizes as indicated. Red arrow indicates estimated molecular weight of H1–3, 38 kDa. **B.** *Left*–Elution profile from 5 mL Reverse Hi–Trap (GE Healthcare) column after TEV cleavage. *Right*–12.5% SDS–PAGE analysis of eluted fractions, of which the flowthrough and fractions 1–3 were pooled. **C.** *Left*–Elution profile from 300 mL Superdex 75 (GE Healthcare) column. *Right*–12.5% SDS–PAGE analysis of eluted fractions, of which fractions C2–C5 were pooled.

removing the HUBL domains significantly decreases the activity as observed with $USP7_{CD}$ and TRAF–$USP7_{CD}$.

Table 3.7
Purification Summary and Molar Extinction Coefficients of USP7 Truncations

Truncation	Molecular Weight (Da)	Extinction Coefficient ($M^{-1}*cm^{-1}$)	Yield (mg)	Specific Activity (Units*mg^{-1})
$USP7_{FL}$	125,710	125,584*	5	0.1
$USP7_{CD}$	41,265	43,316	22	0.003
$USP7_{CD}$–H4–5	67,811	56,893	1.5	0.2
$USP7_{CD}$–H1–5	104,405	89,579*	3.6	0.2
TRAF–$USP7_{CD}$	66,177	74,250*	1.2	0.002
H1–5	63,665	52,842	5.6	N/A
H4–5	25,329	21,403	6	N/A
H1–2	24,801	25,495	44	N/A
H1–3	37,737	33,284	31	N/A

* Values were determined computationally with Expasy due to low concentration or purity [52]. Unit is defined as μmol*min^{-1} Yield is per liter of culture

3.5 Discussion

USP7 is a high profile target for the development of anti–cancer therapeutics due to its well characterized role in many disease states. Furthermore, structural and kinetic studies have begun to scratch the surface of the importance of specific interactions between the seven different domains of USP7. However, in our effort to develop a specific USP7 inhibitor, further elucidation of intramolecular interactions of the seven domains of USP7 was essential. Towards this goal, we expressed and purified nine different constructs of USP7. Each construct was codon–optimized, synthesized with a 6His tag and TEV cleavage site, and cloned into the pET–11a expression vector by BioBasic Inc to achieve optimal expression in *E. coli* BL21–DE3 cells. The

constructs expressed within the soluble fraction in *E. coli* BL21–DE3 cells and the purifications utilized the same general sequence of chromatographic steps of nickel affinity, reverse nickel affinity, and size–exclusion chromatography. Only $USP7_{CD}$–H1–5 truncation deviated from this sequence to include a MonoQ chromatographic step in an attempt to improve overall purity of the protein. The purification protocols described in this chapter resulted in milligram quantities of highly pure protein for use with further kinetic and inhibitor characterization. Many of the truncations described here were expressed and purified in an alternative manner by Faesen and coworkers [20] however protein yields were not disclosed within the manuscript and a comparison between the two methodologies cannot be made.

CHAPTER 4. A TETHERED–RHEOSTAT MODEL FOR THE INTRAMOLECULAR ACTIVATION OF THE ANTI–CANCER TARGET USP7/HAUSP BY ITS HAUSP UBIQUITIN–LIKE (HUBL) DOMAINS

4.1 Abstract

Ubiquitin Specific Protease 7 (USP7/HAUSP) is a well–studied deubiquitinating enzyme that regulates many critical signaling pathways in the cell, including regulation of the tumor suppressor p53/MDM2 pathway. The isolated catalytic domain of USP7 has 30–fold lower activity than the full–length enzyme. Previous structural and kinetic studies of USP7 have attributed this lack of activity to the catalytic triad being misaligned. Additional structures have shown that to achieve full activity, the USP7/HAUSP Ubiquitin-like (HUBL) domains 4 and 5 must be present. Previous reports have also suggested that HUBL domains 1–3 are simply scaffolds for protein–protein interactions with USP7 binding partners, such as the allosteric activator GMPS. Here, the kinetic differences between the activation of the isolated catalytic domain of USP7 ($USP7_{CD}$) *in trans* by HUBL domains 1–5 and activation by HUBL domains 4–5 were evaluated, and kinetic and biochemical characterization of these observed differences revealed the importance of HUBL domains 1–3 in the intramolecular activation of USP7. A new model for USP7 regulation is proposed where by the HUBL domains 1–3 act as a rheostat, regulating the level of activation USP7 experiences both internally as well as by allosteric activator GMPS.

4.2 Introduction

One of the most well studied DUBs is the ubiquitin specific protease 7 (USP7), also known as the herpes virus-associated ubiquitin-specific protease (HAUSP). USP7

is renowned for its essential role in the deubiquitination of both the tumor suppressor protein p53 and its E3 ligase, MDM2 [40–42, 56]. When the cell is under normal conditions, USP7 primarily deubiquitinates MDM2. However, when the cell experiences genotoxic stress, USP7 shifts its deubiquitination from MDM2 to p53 [33, 57]. Furthermore, under genotoxic stress, guanosine 5–monophosphate synthase (GMPS) binds to and allosterically activates USP7 to enhance the rate of deubiquitination which leads to the stabilization of p53 [58].

Disruption of the p53 signaling pathway is responsible for roughly 50% of human cancers due to mutation or inactivation of p53 [38]. By deubiquitinating MDM2 and p53, as well as other known USP7 substrates such as tumor suppressors PTEN and FOXO4, and viral E3 ligase ICP0, USP7 has a hand in pathways involved in viral infection, cell signaling, proliferation, and apoptosis [35–37, 39, 59]. Therefore, tight regulation of USP7 activity is essential to maintaining proper cell signaling, and the up–regulation of USP7 is a major contributing factor in the progression of many cancers [35, 36, 40–43, 56]. Two research groups have identified compounds that show potential for the specific inhibition of USP7 [45, 60]. For example, Chauhan and coworkers have identified a compound series that inhibit cell growth and induces apoptosis in multiple myeloma cell lines [45]. Alternatively, Colland and coworkers have identified a compound that can increase cellular levels of p53 leading to p53–mediated apoptosis [60].

X–ray crystal structures for each domain of USP7 are available as individual proteins, however, a full–length structure has yet to be determined. USP7 is comprised of seven domains (Figure 4.1 A); a TNF receptor associated factor (TRAF) domain, a catalytic domain, and five ubiquitin-like (UBL) domains [19, 20, 33]. The N–terminal TRAF domain is proposed to be responsible for substrate binding and specificity [33]. The catalytic domain maintains the conserved papain–like fold seen throughout the USP family [19]. Interestingly, the purified catalytic domain is 120–fold less active than that of the full–length enzyme as measured with the fluorogenic substrate ubiquitin–7-amino–4-methylcoumarin (Ub–AMC) [34]. The poor activity

of the catalytic domain has been attributed to the observation that the catalytic triad is misaligned in the unbound structure of USP7 (PDB: INB8 and 4M5W) presumably making it catalytically inefficient (Figure 4.1 B) [15, 16]. However, in a crystal structure of the catalytic domain bound to ubiquitin aldehyde (PDB: 1NBF), the catalytic triad is properly aligned for catalysis, suggesting that an induced–fit rearrangement of the catalytic triad by binding of the substrate may be required for catalysis (Figure 4.1 B) [16]. Furthermore, a comparison of the unbound structure of the catalytic domain to the ubiquitin aldehyde–bound structure shows that significant movement of blocking loops 1 (residues 408-429, BL1) and 2 (residues 459-462, BL2) as well as the switching loop (residues 285-291) is required to unblock the active site for the binding of ubiquitin–aldehyde (Figure 4.1 C and D) [15, 19].

At the C–terminus of USP7 are five consecutive UBL domains referred to as the HAUSP ubiquitin-like (HUBL) domains [20]. UBL domains are found throughout nature and are surprisingly commonplace in USPs [22, 23]. There are sixteen USPs that have been described to have UBL domains [22, 23]. UBL domains are defined by the fact that they contain the highly stable β–grasp fold adopted by ubiquitin. The sequence identity of UBL domains relative to each other and to ubiquitin is very low, and, as a result, UBL domains are often only identified by determining their structures [22]. This remains true for the HUBL domains where the sequence identity between them is only 3 to 15%, or to ubiquitin, varying from 6 to 19% [20]. Unlike ubiquitin, UBLs do not contain the two terminal glycine residues and cannot be conjugated to protein substrates.

Information regarding the function of the UBL domains and their interactions within their respective proteins is lacking [22]. One UBL that is beginning to be understood is the one within USP14. A UBL domain is present in the structure of USP14, and is responsible for the association of USP14 with the proteasome [19]. The binding of USP14 to the proteasome opens up the active site that is blocked by BL1 and BL2. The net result of the unblocking is an increase in catalytic activity [19, 61]. In 2005, the Mesecar lab identified the first viral UBL domain in the Severe Acute

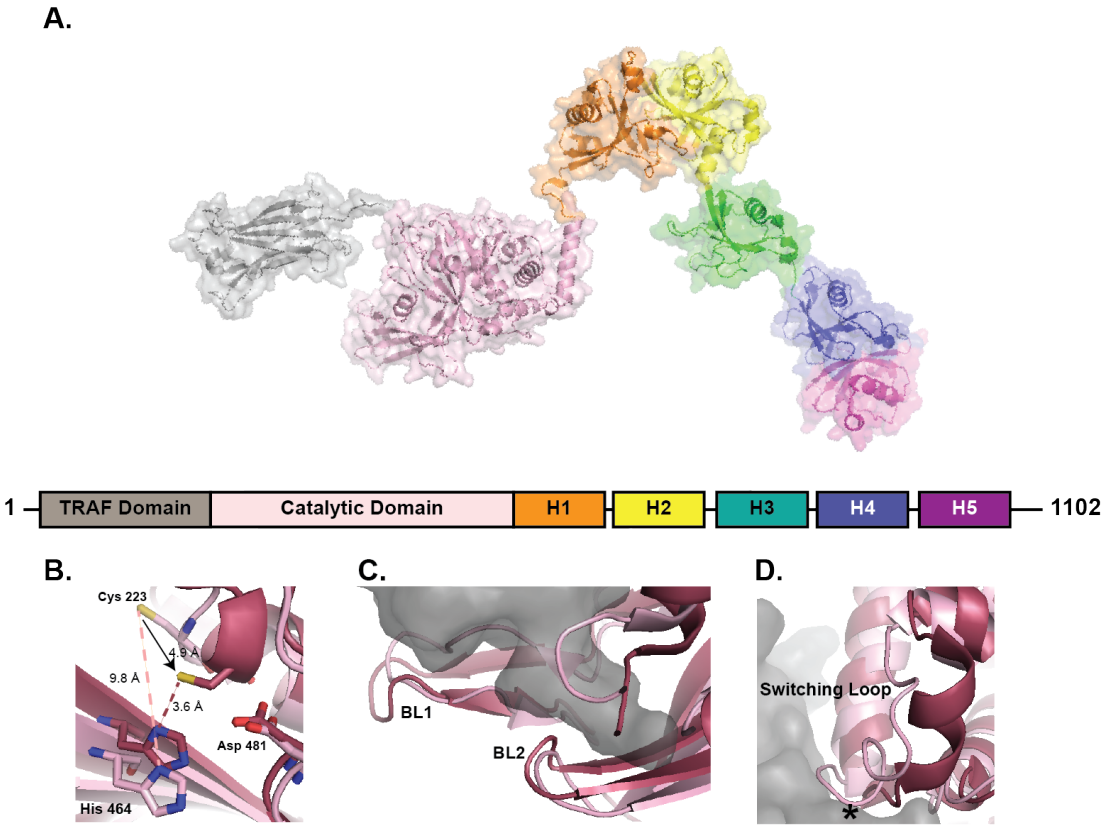


Fig. 4.1. **Domain Organization of USP7 and Active Site Rearrangement Upon Ubiquitin Binding A.** Structural representation of the X–ray crystal structures of the seven domains of USP7. Grey: TRAF domain (PDB: 2F1W). Light pink: catalytic domain (PDB: 4M5W). Orange: HUBL1, Yellow: HUBL2, Green: HUBL3, Blue: HUBL4, Purple: HUBL5 (PDB: 2YLM). **B.** A shift in the catalytic triad occurs between unbound (PDB: 4M5W) and ubiquitin–aldehyde bound catalytic domain (PDB: 1NBF). The distances between the cysteine and the histidine in each form of the enzyme, as well as how far the cysteine shifts, is labeled. **C.** Movement of the blocking loops BL1 and BL2 must occur to accommodate ubiquitin binding. **D.** A shift in the switching loop (residues 285-291) in the ubiquitin–aldehyde bound structure alleviates steric hindrance (denoted by the asterisk).

Respiratory Syndrome virus papain–like protease (SARS-PLpro). This single UBL domain is thought to be involved in the evasion of the virus from the innate immune system but its definitive function is still unknown [62]. While USP14 and SARS–PLpro only contain one UBL domain, it is common for USPs to have multiple UBL domains [22]. However, no USP has been previously described to have five tandem UBLs, a feature that is so far unique to USP7. Thus understanding the roles of the five HUBL domains in USP7 function is of scientific interest.

A number of X–ray crystal structures of the HUBL domains has been determined with different constructs including the complete HUBL domain (PDB 2YLM), HUBL domains 1–3 in complex with peptide fragments of USP7s substrates ICP0 and DNMT1 (PDB: 4WPH, 4WPI, and 4YOC), as well as a truncation of USP7 containing just the catalytic domain and HUBL domains 1–3 (Kim et al, and PDB: 5J7T) [20, 37, 39, 63, 64]. The complete structure of the HUBL domains suggests that the individual domains are arranged in a 2–1–2 fashion and can assume an elongated orientation (Figure 4.1, PDB 2YLM) [20, 37]. Two X–ray crystal structures of HUBL domains 1–3 were determined by Pfoh and coworkers and the structures show the HUBL domains in two different orientations. These two structures provide insight into the flexibility of these domains which occurs through a hinge between HUBL domains 2 and 3 (2/3 hinge) which allows the HUBL domains to assume a compact orientation in addition to the elongated structure [39]. The X–ray crystal structure of the catalytic domain with HUBL domains 1–3 was recently solved by Kim and coworkers. The structure reveals that the HUBL domains are in a compact orientation, supporting the identification of the 2/3 hinge. The structure also revealed the precise location of the α–helix or "tether" between the catalytic domain and the HUBL domains [63]. Biochemical and mutagenesis studies of the rigidity, length, and localization of charged groups in the tether elucidated its role in the spatial orientation of the HUBL domains with respect to the catalytic domain [63].

It has long been understood that the HUBL domains are required for full activity of USP7. A kinetics study by Faesen and coworkers suggested that only HUBL

domains 4 and 5 are truly required for full activation of USP7 [20, 34]. This finding was supported by X–ray crystal structures determined by Rougé and coworkers (PDB: 5JTV and 5JTJ) which show the C–terminal peptide of HUBL 5 binding in an activation cleft, a pocket apparent in X–ray crystal structures of the catalytic domain bound to ubiquitin [64]. Kinetic studies by both groups support the importance of this peptide in the activation of USP7 [20, 64]. Rogué and coworkers have hypothesized that the binding of this peptide stabilizes the ubiquitin–USP7 complex for increased catalytic activity [64]. However, a detailed mechanism of how this interaction influences USP7s activity has yet to be fully elucidated [20]. The activation USP7 experiences by the binding of this loop within the activation cleft, however, can be easily disrupted by a single point mutation [20, 64]. Furthermore, it has been suggested that HUBL domains 1–3 are not involved in the activation process and are dispensable for intramolecular activation [20]. However, an acidic binding pocket for allosteric activator GMPS and substrate DNMT1 was discovered in HUBL domain 2 which may shed light on the importance of HUBL domains 1–3 [20, 37, 39].

In this dissertation work, the kinetic mechanism by which the HUBL domains influence the activity of USP7 is explored and the kinetic data support of the structural model presented by Rougé and coworkers [64]. The results also support some of the previous kinetic findings that all five HUBL domains, as well as just the HUBL domains 4 and 5, can increase the activity of USP7. However, through an in depth kinetic analysis of the role of the HUBL domains, significant differences in the kinetic parameters that influence the activity of USP7 are observed, and these differences depend upon which HUBL domains are present. Previously, the role of HUBL domains 1–3 was described as a binding site for substrates as well as the allosteric activator GMPS, and did not have a direct role in the intramolecular activation of USP7 [20, 39]. The new findings presented here shed light on the requirement of HUBL domains 1–3, and their influence on HUBL domains 4 and 5 to achieve the native form of intramolecular activation of USP7.

4.3 Experimental Procedures

4.3.1 Kinetic Assays

Steady–State Kinetic Studies

The K_m and k_{cat} values of $USP7_{FL}$ and $USP7_{CD}$ were determined with Ub–Rho110. The purified enzymes were diluted to a working concentration of 2 nM in the Assay Buffer described in Section 2.3.12. Twelve, two–fold serial dilutions of Ub–Rho110 were made ranging from 0.04 μM to 40 μM in Substrate Buffer described in Section 2.3.12. Assays were performed in 30 μL reaction volumes and in triplicate. The background rate were measured by pipetting 15 μL into a 96–well Corning Costar black half–volume plate and measuring the rate of change in fluorescence with a BioTek Synergy H1 plate reader. Fluorescence measurements were made every 10 seconds at an excitation λ of 360 nm and emission λ of 460 nm over a time period of two minutes and at 25 °C. Following the measurement of the background, 15 μL of the working stock of the USP7 enzymes was added to the plate, the plate was mixed for 10 seconds, and the assay was measured under the same conditions for 30 minutes total, or until the reactions started to plateau. The final concentration of the USP7 enzymes were 1 nM and the final concentration range for the substrates was 0.01 μM to 20 μM. The initial slope of each reaction in arbitrary fluorescence units (AFU) per unit time (AFU/sec) was converted to the amount of product produced per unit time (M/sec) with the extinction coefficient of the substrate. The extinction coefficient of both substrates was determined as described for Ub–Rho110 in Section 2.3.12. This rate value in μM/sec was then converted to the turnover number (V/[E]) by dividing by the enzyme concentration. The data were then analyzed with the kinetics module of SigmaPlot (v13: Systat Software Inc.) by fitting the resulting curves to the Michealis–Menten equation (Equation (2.1)) to determine the k_{cat} and K_m for each enzyme.

Kinetic Studies on the Influence of the HUBL Domains on the Interaction of Ub–Rho110 with USP7

The kinetic influence the HUBL domains have on the activity of USP7 was kinetically evaluated with $USP7_{CD}$ and HUBL truncations H1–5 and H4–5 *in trans*. The kinetic assays were performed by varying the concentration of Ub–Rho110 at a fixed concentration of the HUBL truncation and vice versa. Two–fold serial dilutions of Ub–Rho110 were made in Substrate Buffer to obtain a working concentration range of 0.12 μM to 120 μM. Two–fold serial dilutions of HUBL truncations H1–5 and H4–5 were made in Assay Buffer to obtain a working concentration range of 0.075 μM to 37.5 μM. $USP7_{CD}$ was diluted in assay buffer to a working concentration of 8.6 nM. Each assay was performed in a 30 μL reaction volumes in triplicate in a Corning 96 half–well black plates. The background rate were measured by pipetting 5 μL of Ub–Rho110 and 20 μL of the HUBL truncation into a 96–well Corning Costar black half–volume plate, and measuring the rate of change in fluorescence with a BioTek Synergy H1 plate reader. Fluorescence measurements were made every 10 seconds at an excitation λ of 360 nm and emission λ of 460 nm over a time period of two minutes and at 25 °C. To initiate the reaction, 5 μL of $USP7_{CD}$ was added to the wells and the plate was read for 30 minutes. The final concentrations of each component is as follows: Ub–Rho110 ranged from 0.02 μM to 20 μM, HUBL truncation ranged from 0 μM to 25 μM and the $USP7_{CD}$ concentration was held constant at 1 nM. The initial rates were determined and converted to v/[E] using the extinction coefficient of the substrate as described in Section 2.3.12. The resulting v/[E] values were plotted as a function of Ub–Rho110 or as a function of HUBL truncation. The curves of v/[E] as a function of substrate were fit to the activation models provided in the kinetics module of SigmaPlot (v13: Systat Software Inc.) to statistically determine the best model for this system. The equation describing a kinetic model for the non–essential activation (Figure 4.2) was used to fit the data,

$$\frac{v}{V_{\max}} = \frac{[S]}{K_m \frac{\left(1+\frac{[A]}{K_{A,app}}\right)}{\left(1+\frac{\beta[A]}{\alpha K_{A,app}}\right)} + [S] \frac{\left(1+\frac{[A]}{\alpha K_{A,app}}\right)}{\left(1+\frac{\beta[A]}{\alpha K_{A,app}}\right)}} \tag{4.1}$$

where E represents $\mathrm{USP7_{CD}}$, A is a HUBL truncation (H1–5 or H4–5), and S is Ub–Rho110. For the non–essential activation model, enzyme velocity is given by

$$v = k_{cat}[ES] + \beta k_{cat,app}[EAS] \tag{4.2}$$

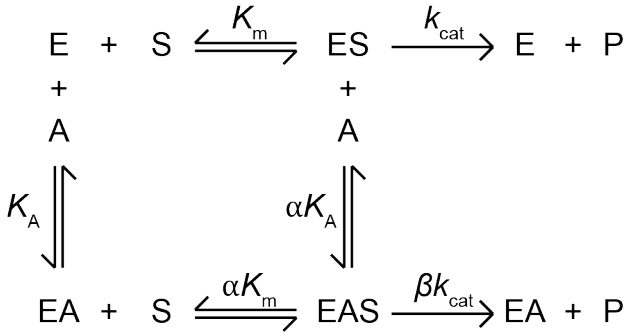


Fig. 4.2. **Non–essential Kinetic Model.** Expansion of the simple enzyme kinetic scheme to include the potential effects of an activator and the introduction of $\mathrm{K_A}$ kinetic parameter and kinetic modulators α and β which manipulate $\mathrm{K_A}$, $\mathrm{K_m}$, and k_{cat}. **E**: enzyme, $\mathrm{USP7_{CD}}$, **A**: activator, HUBL truncation, **S**: substrate, Ub–Rho110

From the fit to Equation (4.1), the kinetic parameters: $K_{m,app}$, $k_{cat,app}$, $K_{A,app}$, and the kinetic modifiers α, and β in the presence of a particular HUBL truncation were determined. As described in Section 4.4.2 the data were also fit to the substrate–activator model as well, however the non–essential activation model was determined to be the best fit by comparing the AICc values.

4.3.2 Di–Ubiquitin Chain Cleavage Assays

Evaluation of the ability of a particular HUBL truncation to activate $USP7_{CD}$ with the more native substrate Lys^{48} was performed with a reaction mixture composed of 4.2 μM of Lys^{48} di–ubiquitin (Boston Biochem), 50 nM $USP7_{CD}$, and either 2, 6, or 20 μM of H1–5 or H4–5. A negative control was included with no enzyme. Control of $USP7_{FL}$ and $USP7_{CD}$ with Lys^{48} was also included. All reactions were performed at 25 °C in Assay Buffer for two hours. Reactions were quenched with the addition of NuPAGE LDS sample buffer (Invitrogen) supplemented with 5 mM fresh DTT. Samples were analyzed by SDS-PAGE using NuPAGE Novex 4–12% BisTris mini gels (Invitrogen) stained with Coomassie Brilliant Blue.

4.3.3 Isothermal Calorimetry (ITC) Analysis of Compound Binding

Isothermal calorimetry was utilized to further evaluate the interaction between $USP7_{CD}$ and the HUBL domain truncations: H1–2, H4–5, and H1–5. All proteins were dialyzed overnight at 4 °C into Analytical Buffer (50 mM HEPES pH 7.5, 3 mM TCEP, and 100 mM NaCl). $USP7_{CD}$ was concentrated by ultrafiltration with a 10,000 MWCO membrane (Millipore) and diluted to 30 μM in Analytical Buffer. HUBL truncations were concentrated in the same fashion with the appropriate MWCO membrane. Each truncation was diluted to 1.5 mM in Analytical Buffer (50 fold higher than [$USP7_{CD}$]). A Malvern MicroCal ITC200 instrument was used for the measurements. A volume of 200 μL of 30 μM $USP7_{CD}$ was added to the cell, and the syringe volume of the HUBL truncation was 100 μL. The initial injection volume was 0.4 μL to remove any bubbles from within the system, and the remaining injections had a volume of 1.8 μL for each USP7 truncation and a total of 22 injections was performed (40 μL). A three minute delay was programmed between each injection to allow the signal to return to baseline. All reactions were run at 25 °C. The spin speed was set to 1000 rpm. For each truncation, blank injections were run by injecting the USP7 truncation into Analytical Buffer. Data was exported into NITPICK software (NIH)

and the heat from the blank run was subtracted, and the results were analyzed for measurable heat.

4.3.4 Interference of $USP7_{FL}$ Activity or $USP7_{CD}$–H4–5 Activity by HUBL Truncations *in trans*

To determine the amount of interference caused by the individual HUBL truncations *in trans* with $USP7_{FL}$, the total concentration of Ub–Rho110 was fixed at 1.25 μM, and the $USP7_{FL}$ concentration was fixed at 1 nM. Each HUBL truncation tested (H1–3, H1–2, H1–5, and H4–5) was prepared as a 1.3x working stock (260 μM). Two–fold serial dilutions of the HUBL truncations were made resulting in 11 concentrations of each HUBL truncation ranging from 0.2 μM to 200 μM. A 10x working stock of $USP7_{FL}$ and 6x working stock of Ub–Rho110 were prepared. All proteins were diluted in Assay Buffer and Ub-Rho110 was diluted in Substrate Buffer. To a 96–well Corning Costar black half–volume plate, 5 μL of the working stock of Ub–Rho110 and 23 μL of each concentration of the particular HUBL truncation being tested was added to individual wells. The plate was mixed for 10 seconds within the plate reader and the background rates were measured for two minutes at conditions described previously (Section 2.3.12). To initiate the reaction, 3 μL of the working stock of $USP7_{FL}$ was added to the plate, and the plate was shaken within the plate reader for 10 seconds prior to measuring the activity. The enzymatic activity was measured for 30 minutes as described previously (Section 2.3.12). The enzymatic activity of $USP7_{FL}$ alone was measured also measured for comparison. All assays were done in triplicate. The % Interference was calculated for each concentration of the HUBL truncations using Equation (4.3).

$$\%\ \text{Interference} = \left(1 - \frac{\text{Rate}_{\text{sample}} - \text{Rate}_{\text{bkgd}}}{\text{Rate}_{\text{no interference}} - \text{Rate}_{\text{bkgd}}}\right) \times 100 \tag{4.3}$$

In Equation (4.3), $Rate_{sample}$ is the initial slope of the original progress curve of $USP7_{FL}$ in units of AFU/second in the presence of a particular HUBL truncation and concentration. The $Rate_{nointerference}$ is the initial slope of the progress curve of $USP7_{FL}$ without any interfering HUBL truncation. The $Rate_{bkgd}$ is the background rate of the Ub-Rho110 in the presence of just the HUBL truncations in the absence of $USP7_{FL}$. This control allows us to confirm that no HUBL truncation has enzymatic activity itself. The calculated % Interference were then plotted in terms of HUBL truncation concentration. The resulting curve was fit to the Michealis–Menten equation (Equation (2.1)) with SigmaPlot as described previously (Section 2.3.12). The calculated K_m value from the fit corresponds to the IC_{50} values for the HUBL truncation, and the V_{max} value corresponds to the maximum % Interference inflicted on the system by the HUBL truncation (Table 4.3). The same experimental set–up and data analysis was utilized to evaluate the interference of H1–3 and H4–5 with $USP7_{CD}$–H4–5.

4.3.5 Analytical Size–Exclusion Chromatography to Evaluate Any Direct Interaction Between H1–3 and H4–5

USP7 truncations H1–3 and H4–5 were evaluated for direct interaction by size exclusion chromatography (SEC). H1–3 was diluted to a final concentration of 65 μM in Analytical Buffer (50 mM Tris pH 7.5, 3 mM TCEP, and 100 mM NaCl) in a final volume of 200 μL. H1–3 was then injected onto a 24 mL Sepharose 6 (GE Healthcare) SEC column and run at a flow rate of 0.5 mL/min. After the run of H1–3 alone was complete, H4–5 was diluted to 65 μM in 200 μL and run identically. To determine if H1–3 and H4–5 have a direct interaction, a 1:1 molar ratio of H1–3 to H4–5 was mixed at 65 μM of each component in a volume of 200 μL. The presence of an interaction was evaluated by the elution profile of the combined H1–3 and H4–5 chromatogram compared to the individual chromatograms of H1–3 and H4–5.

4.4 Results

4.4.1 Steady–state kinetic parameters of $USP7_{FL}$, $USP7_{CD}$, and $USP7_{CD} - H4 - 5$

The full–length USP7 (residues 1-1102, $USP7_{FL}$) and the catalytic domain (residues 206-560, $USP7_{CD}$) were purified to > 98% purity from *Sf9* cells and *E. coli* cells, respectively (Chapter 3). The steady–state kinetic characterization of the deubiquitinating activity of each enzyme was performed with the fluorogenic substrate Ub–Rho110. The initial rates were measured in triplicate and plotted versus increasing Ub–Rho110 concentrations (Figure 4.3A and B). The resulting data were fit to the Michealis–Menten equation (Equation (2.1)), and the K_m and k_{cat} values derived from the best–fits for each enzyme combination are summarized in Table 4.1. The resulting K_m value for $USP7_{FL}$ was 1.46 ± 0.07 μM and the k_{cat} value was 1.24 ± 0.02 s^{-1}. The K_m value of $USP7_{CD}$ at 1.5 ± 0.5 μM is the same within error as that of $USP7_{FL}$. However, the k_{cat} value of $USP7_{CD}$ is only 0.040 ± 0.002 s^{-1} which is a 30–fold decrease in the maximal rate of $USP7_{CD}$ compared to $USP7_{FL}$. When the steady–state kinetic parameters were previously determined using Ub–AMC as the substrate, a 120–fold difference in the maximal rate between $USP7_{FL}$ and $USP7_{CD}$ was observed (Table 4.1) [20]. Moreover, the K_m value of $USP7_{CD}$ for Ub–Rho110 is lower than that for Ub–AMC (15.13 ± 1.09 μM). Ub–Rho110 was therefore chosen for kinetic studies as the concentration of Ub–Rho110 required to reach saturation relative to Ub–AMC is much lower which will lower the cost substantially, reduce the chance of inter–filter effects, and allows the kinetic studies to be run over a larger range of concentrations providing greater accuracy.

The catalytic domain of USP7 tethered to HUBL domains 4 and 5 (residues 206-1102 Δ564-879, $USP7_{CD}$–H4–5) was purified to > 98% purity from *E. coli* as determined by SDS–PAGE analysis (Chapter 3). The steady–state kinetic parameters were determined with Ub–Rho110 as described previously (Figure 4.3C). The resulting K_m for $USP7_{CD}$–H4–5 was 2.2 ± 0.2 μM and the k_{cat} was 1.33 ± 0.05 s^{-1} (Table 4.1).

With Ub–AMC the reported K_m and k_{cat} values with $USP7_{CD}$–H4–5 were 2.89 ± 0.20 μM and 1.32 ± 0.03 s^{-1}, respectively [20]. These results support those by Faesen and coworkers who found that HUBL domains 4 and 5 can activate USP7 to a similar extent as with all HUBL domains present in $USP7_{FL}$ [20].

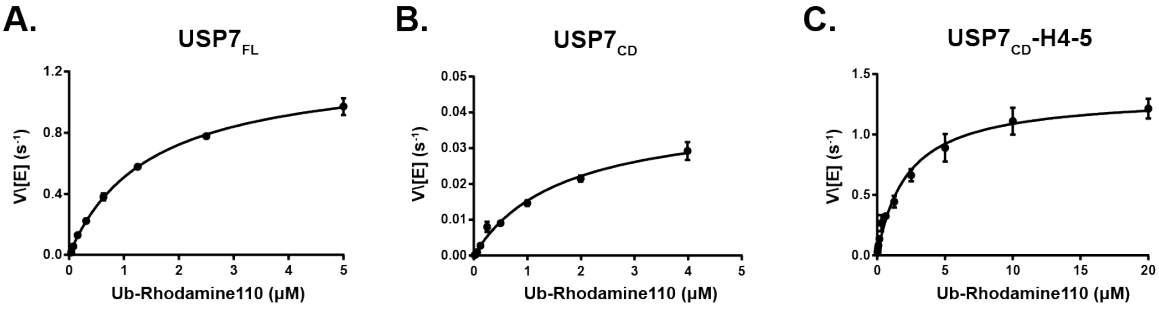


Fig. 4.3. **Steady–State Kinetic Response of $USP7_{FL}$, $USP7_{CD}$, and $USP7_{CD} - H4 - 5$ to the Substrate, Ub–Rho110.** Rates for the enzymatic activity are normalized to the total enzyme concentration and plotted against varying Ub–Rhodamine110 concentration. **A.** $USP7_{FL}$ at 1 nM, **B.** $USP7_{CD}$ at 200 nM, and **C.** $USP7_{CD}$–H4–5 at 1 nM. Ub–Rhodamine 110 concentrations varied from 0.02 μM to 20 μM depending on the enzyme being tested. Error bars, Standard Deviation.

4.4.2 HUBL domains 1–3 modulate the steady–state kinetic parameters surrounding activation of USP_{CD} *in trans*

Previous work has suggested that HUBL domains 4 and 5 (H4–5) are required for activating $USP7_{CD}$ [20,34]. It was further suggested that HUBL domains 1–3 (H1–3) do not play a role in the intrinsic USP7 activity, and are only required for binding to the allosteric activator GMPS which increases the k_{cat} over 5–fold [20]. A kinetic approach was taken to further evaluated the role of H1–3 in the intramolecular activation of USP7. Activity assays were performed with $USP7_{CD}$ by varying the concentrations of either H1–5 or H4–5 *in trans* as well as, using fixed, variable concentrations of Ub–Rho110. The initial rates of hydrolysis were measured for each condition, and plotted with respect to either the concentration of Ub–Rho110 (Figure 4.4A and B) or to the

Table 4.1
Steady–state kinetic parameters of $USP7_{FL}$, $USP7_{CD}$, and $USP7_{CD}$–H4–5 constructs as measured with fluorogenic ubiquitin substrates

	Ub–Rho110			**Ub–AMC**[a]		
	$USP7_{FL}$	$USP7_{CD}$	$USP7_{CD}$–H4–5	$USP7_{FL}$	$USP7_{CD}$	$USP7_{CD}$–H4–5
K_m (μM)	1.46 ± 0.07	1.5 ± 0.5	2.2 ± 0.2	2.89 ± 0.10	15.13 ± 1.09	2.89 ± 0.2
k_{cat} (s^{-1})	1.24 ± 0.02	0.040 ± 0.002	1.33 ± 0.05	1.42 ± 0.02	0.06 ± 0.02	1.32 ± 0.03
k_{cat}/K_m ($x10^6 * s^{-1} * M^{-1}$)	0.85 ± 0.04	0.012 ± 0.004	0.60 ± 0.07	0.49	0.004	0.46

[a] Values reported by Faesen and coworkers [20].

concentration of H1–5 or H4–5 HUBL truncation, (Figure 4.4C and D). All conditions produced hyperbolic saturation curves for the response of $USP7_{CD}$ to the binding of the Ub–Rho110 substrate. Surprisingly, distinct differences were observed between activation by H1–5 and H4–5 *in trans*. At 25 μM of H1–5, the activity of $USP7_{CD}$ is increased 40–fold relative to $USP7_{CD}$ alone, while with H4–5, the activity of $USP7_{CD}$ is increased only 20–fold at 25 μM H4–5 (Figure 4.4A and B). Figure 4.4D shows that H4–5 just starts to reach saturation at 25 μM with Ub–Rho110, while Figure 4.4C shows that H1–5 fully saturates within the concentration range tested (0 μM to 25 μM).

The HUBL domains have been described by Faesen and coworkers to activate USP7, so the data in Figure 4.4 were fit to the non–essential activation model (Equation (4.1)) and the substrate–activator (SA) model using the Enzyme Kinetics Module of SigmaPlot (v13: Systat Software Inc.) [20]. The SA model requires the substrate to first interact with the activator to form the true substrate of the enzyme prior to binding to the enzyme [65]. This is in contrast to the non–essential activation model which assumes that the reaction can occur in the absence of the activator all together, just at a reduced rate [65]. It is unlikely that the HUBL domains activate by substrate-activator (SA) activation, as $USP7_{CD}$ has activity without the HUBL domains present. Rather, it is more likely that USP7 is an example of non–essential activation. To test this hypothesis, the Akaike Information Criterion for small sample size (AICc) values were compared from the fits to each equation. The best model is determined by the fit that resulted in the most negative AICc value, as the more negative the AICc value is, the better the fit. It was clear that for both the H1–5 and H4–5 data sets, the fit to the non–essential activation model was best. For H1–5, the AICc value for the fit to the non–essential activation model was -385, which is 167 points more negative than that of the SA model (-218). This trend is observed for H4–5 as well; the AICc value for the non–essential activation model (-689) was 312 points more negative than that of the SA model (-377). Using the equation for non–essential activation (Equation (4.1)) the kinetic parameters K_m, k_{cat}, K_A, α, and

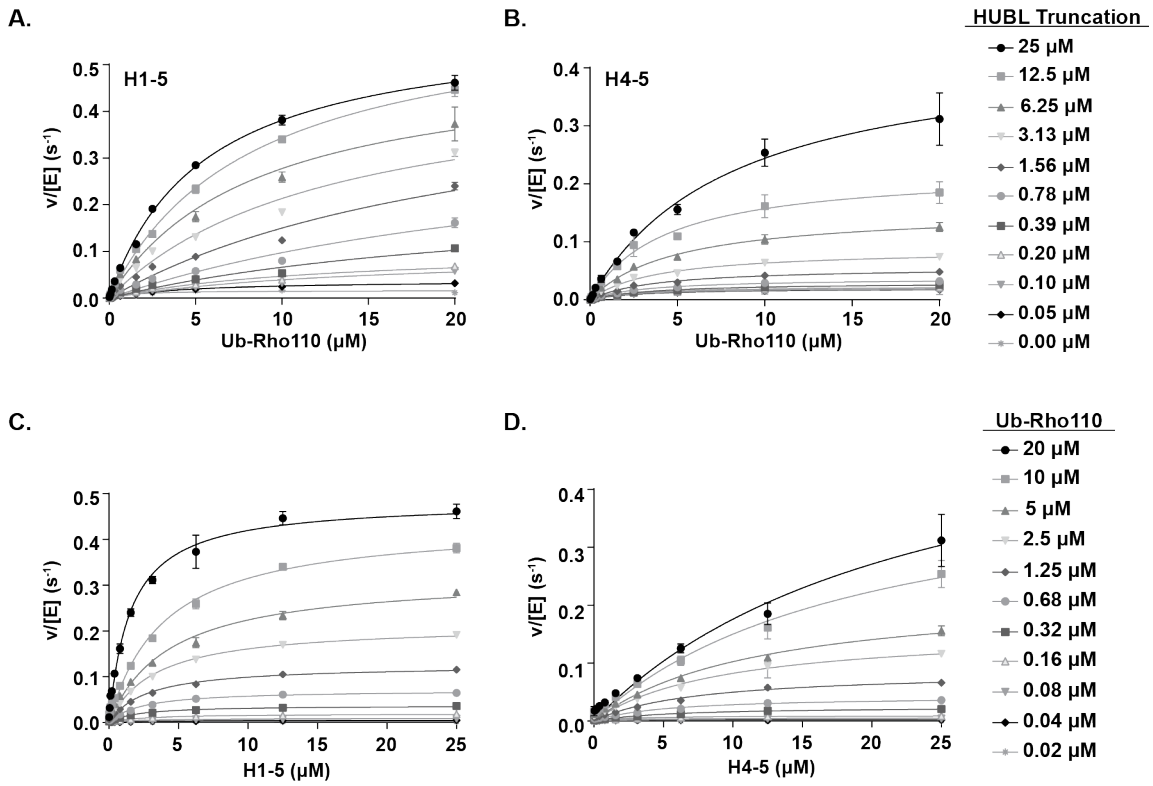


Fig. 4.4. **Steady–State Kinetic Analysis of USP7$_{CD}$ with HUBL Domains *in trans*.** A survey of steady–state kinetic parameters of the interaction of USP7$_{CD}$ with HUBL truncations H1–5 or H4–5 *in trans*. Reactions were performed by varying the concentration of Ub–Rho110 from 2–20 μM (A and B) or HUBL truncation concentration from 0–25 μM (C and D) with 1 nM USP7$_{CD}$. Initial AFU/s were measured for each curve, and were converted using the extinction coefficient of the substrate to v/[E]. Resulting values were plotted as v/[E] in terms of the concentration of the constant variable. For **A.** and **C.**, H1–5 was tested. Curves **B.** and **D.** are the results of USP7$_{CD}$ with H4–5 *in trans*.

β were determined for each fit (Table 4.2). The α– and β–values are factors by which K_A, K_m, and k_{cat} change when the substrate is engaged with the enzyme. A large difference in the K_m value for $USP7_{CD}$ with H1–5 present (40 $\pm$ 7 μM) is observed in comparison to $USP7_{CD}$ alone (1.5 $\pm$ 0.5 μM) but not with H4–5 and $USP7_{CD}$ (2.6 $\pm$ 0.3 μM). This discrepancy is discussed further in Section 4.4.4.

Interestingly, with no substrate present, the global fit to the non–essential activation model reported that the $K_{a,app}$ value for H1–5 (8.7 $\pm$ 0.7 μM) and H4–5 (9 $\pm$ 1 μM) are identical. This finding is supported by the SPR findings by Faesen and coworkers [20]. They determined the dissociation constant for $USP7_{CD}$ and H1–5 as 48.0 $\pm$ 9.7 μM, however no interaction could be measured for $USP7_{CD}$ and H1–3. They therefore concluded that the HUBL domains associate with $USP7_{CD}$ by H4–5 alone [20].

To further validate the findings of Faesen and coworkers isothermal calorimetry was utilized to measure binding. Three HUBL domains were tested: H1–2, H4–5, and H1–5. The HUBL truncations were loaded into the syringe and injected into $USP7_{CD}$ in the ITC cell. For all combinations tested no heat could be detected above background. This observation suggests one of two possibilities. First, the interaction between $USP7_{CD}$ and the HUBL truncations is not enthalpically driven and no measurable heat is released. The second possibility is that the interactions are too weak to measure. The latter argument is supported by the SPR results of Faesen and coworkers who observed that the interaction between $USP7_{CD}$ and the HUBL domains is weak *in trans*.

H1–5 activation of $USP7_{CD}$ reduces the K_m and K_A values in the presence of substrate with an α–value of 0.11 $\pm$ 0.02 and increases the k_{cat} value with a β–value of 6.2 $\pm$ 0.9. In contrast, α and β for H4–5 are both greater than one (10 $\pm$ 3 and 100 $\pm$ 19, respectively), resulting in an increase in K_m, K_A, and k_{cat} in the presence of substrate (Table 4.2). This difference in kinetic parameters suggests that with H1–3 present, H1–5 modifies the kinetic parameters differently than that of H4–5 alone. H4–5 has a substantial increase in k_{cat}, but is hindered by an increase in K_A and K_m

without H1–3 present. The larger error associated with the α– and β–values for H4–5 may be a result of weaker activation at lower concentrations of H4–5.

4.4.3 Activation of USP7$_{CD}$ by H4–5 *in trans* alone is negatively influenced by increasing substrate concentration

To evaluate the appropriateness of applying the non–essential activation model (Figure 4.2) and equation (Equation (4.1)) to this system the data from each curve displayed in Figure 4.4C and D were fit individually to the Michealis–Menten equation. The $\alpha K_{A,app}$ and $\beta k_{cat,app}$ values were determined from each individual concentration of Ub–Rho110 for both H1–5 and H4–5. These $\beta k_{cat,app}$ and $\alpha K_{A,app}$ values were then plotted against the corresponding Ub–Rho110 concentration of the initial individual curve the values were measured from (Figure 4.5A and B, respectively). The resulting curves of the $\beta k_{cat,app}$ for H1–5 and H4–5 are remarkably similar (Figure 4.5A). The $\beta k_{cat,app}$ curves of Figure 4.5A were then fit again to the Michealis–Menten equation to determine a single $\alpha K_{m,app}$ and $\beta k_{cat,app}$ value which represents the entire data set. The determined values are presented on the right side of Table 4.2. The $\alpha K_{m,app}$ values for H1–5 and H4–5 were determined to be 4.9 $\pm$ 0.4 μM and 16.0 $\pm$ 2.0 μM respectively. The $\beta k_{cat,app}$ for H1–5 and H4–5 were determined to be 0.62 $\pm$ 0.02 s^{-1} and 1.08 $\pm$ 0.09 s^{-1}, respectively. The enzymatic efficiency of an enzyme is defined as how many molecules of substrate can be turned over in a unit of time, $k_{cat,app}/K_{m,app}$ and is a value that can be used to compare the activity of different enzymes. The enzymatic efficiency for H1–5 was determined to be 0.13 $\pm$ 0.01 x 10^6 * M^{-1} * s^{-1} and 0.07 $\pm$ 0.01 x 10^6 * M^{-1} * s^{-1} for H4–5 (Table 4.2). These independently derived kinetic parameters are similar to the values derived from the global fit to the non–essential activation model for both H1–5 and H4–5 which provides further support for the application of the non–essential activation model to this system.

A significant difference between the activation by H1–5 and H4–5 can be observed in Figure 4.5B. The $\alpha K_{A,app}$ value for H1–5 shows an initial drop from 4 $\pm$ 1 μM at

Table 4.2
Steady–State Kinetic Parameters of $USP7_{CD}$ with HUBL Domains *in trans* as Measured with Ub–Rho110

	Non–essential Activation Model Fit		**Individual Fit**	
	$USP7_{CD}$+H1–5	$USP7_{CD}$+H4–5	$USP7_{CD}$+H1–5	$USP7_{CD}$+H4–5
K_m (μM)	40 ± 7	2.6 ± 0.3		
k_{cat} (s^{-1})	0.1 ± 0.01	0.020 ± 0.001		
K_A (μM)	8.7 ± 0.7	9 ± 1		
$\alpha K_{m,app}$ (μM)	4 ± 1	26 ± 8	4.9 ± 0.4	16 ± 2
$\alpha K_{A,app}$ (μM)	1.0 ± 0.2	90 ± 29	2.9 ± 0.3	N/A[a]
$\beta k_{cat,app}$ (s^{-1})	0.6 ± 0.1	2.0 ± 0.4	0.62 ± 0.02	1.08 ± 0.09
α	0.11 ± 0.02	10 ± 3		
β	6.2 ± 0.9	100 ± 19		
k_{cat}/K_m ($x10^6 * s^{-1} * M^{-1}$)	0.15 ± 0.05	0.08 ± 0.03	0.13 ± 0.01	0.07 ± 0.01

[a] Indicates an average value could not be determined.

0.02 μM Ub–Rho110 to 2.0 $\pm$ 0.2 μM at 0.32 μM Ub–Rho110 and averages 2.9 $\pm$ 0.3 μM across the concentration range tested. The $\alpha K_{A,app}$ value from the global fit to the non–essential activation model fit for H1–5 is similar at 1.0 $\pm$ 0.2 μM. The $\alpha K_{A,app}$ value for H4–5, however, varies considerably, ranging from 0.6 $\pm$ 0.2 to 23 $\pm$ 5 μM with increasing Ub–Rho110 concentrations. The $\alpha K_{A,app}$ value for H4–5 derived from the fit to the non–essential activation model is 90 $\pm$ 29 μM. These results suggest that with increasing substrate concentration, the weak interaction of H4–5 with $USP7_{CD}$, as described previously in Section 4.4.2, is exacerbated and suggests H1–3 acts to stabilize H4–5 for activation of $USP7_{CD}$ in the presence of substrate.

4.4.4 $USP7_{CD}$ is activated predominantly by changes in $\beta k_{cat,app}$ with respect to HUBL truncation concentration

To further evaluate the requirement of H1–3 in the activation of $USP7_{CD}$, we evaluated the effects of increasing HUBL truncation concentration on the kinetic activity of $USP7_{CD}$ by fitting the data from each individual curve in Figure 4.4A and B to the Michealis–Menten equation. From these fits, individual values for $\beta k_{cat,app}$ and $\alpha K_{m,app}$ were determined in terms of the HUBL truncation concentration (Figure 4.6A and B, respectively). The results show that substantially less H1–5 is required to achieve maximum enzymatic activity in this system, while H4–5 does not reach maximum enzymatic activity within the concentrations tested (Figure 4.6A). Thus, H1–3 does have a significant role in the enzymatic activity of USP7 and does not simply serve as a scaffold for USP7 binding partners.

The $\alpha K_{m,app}$ values obtained from the Michealis–Menten fit of Figure 4.4A and B were plotted in terms of the respective HUBL truncation concentration (Figure 4.6B). The $\alpha K_{m,app}$ values associated with increasing H4–5 truncation concentration shows a gradual but slight increase from 2 to 8 μM. A spike in $\alpha K_{m,app}$ is observed for H1–5 at low concentrations of H1–5 (0.1 to 3.13 μM), which then plateaus at high concentrations of H1–5 ($>$ 3.13 μM) to $\sim$8 μM. This spike may shed light on the

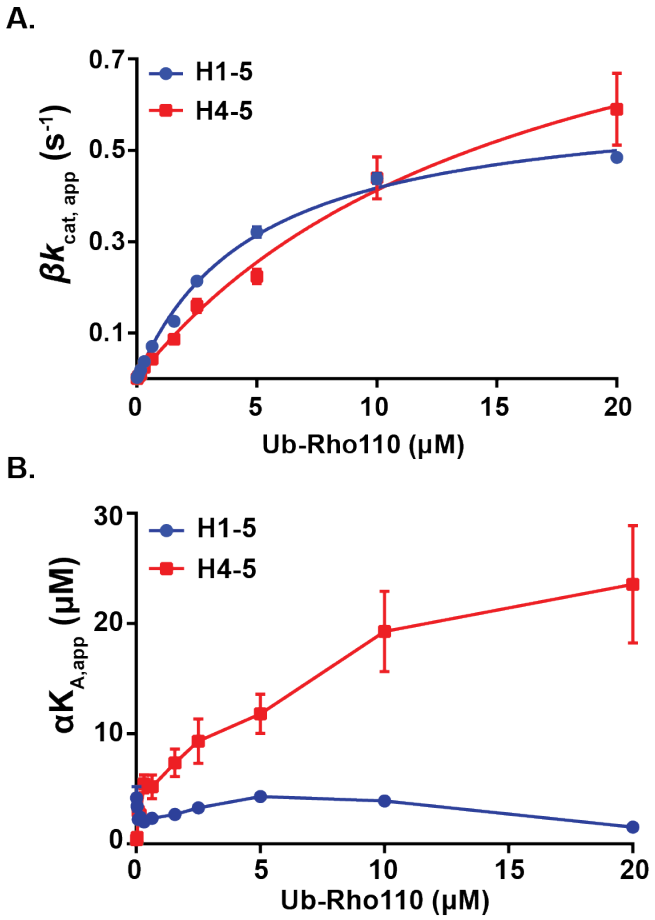


Fig. 4.5. **Modulation of the Steady–State Kinetic Parameters of $\mathbf{USP7_{CD}}$ as a function of Ub–Rho110 Concentration.** The individual curves from Figure 4.4C and D were fit to the Michealis–Menten equation. **A.** The $\beta k_{\mathrm{cat,app}}$ values derived from the fits in terms of Ub–Rho110 concentration. The $\beta k_{\mathrm{cat,app}}$ curves were again fit to the Michealis–Menten equation to derive an overall $\alpha \mathrm{K_{m,app}}$ and $\beta k_{\mathrm{cat,app}}$ value for each truncation. **B.** The $\alpha \mathrm{K_{A,app}}$ values determined from the individual curves in Figure 4.4C and D in terms of increasing Ub–Rho110 concentration. Each assay was done in triplicate. Error bars, Standard Deviation.

observed discrepancy in the global fit of the data to the non–essential activation model K_m value of 40 ± 7 μM for $USP7_{CD}$ with H1–5 *in trans* discussed in Section 4.4.2. To further evaluate the increase in $\alpha K_{m,app}$ observed with H1–5, the $\beta k_{cat,app}/\alpha K_{m,app}$ was determined for each HUBL truncation concentration, and plotted as a function of HUBL truncation concentration (Figure 4.6C). Both H1–5 and H4–5 show a hyperbola with increasing HUBL truncation concentration. The variation in $\alpha K_{m,app}$ for H1–5 does not greatly influence the overall enzymatic efficiency of the system as a function of HUBL concentration, and a mechanistic reason for the spike cannot be further explained at this time. Furthermore, there is only a 2–fold difference between the enzymatic efficiency of $USP7_{CD}$ by H1–5 and H4–5 with respect to HUBL truncation concentration, which supports our previous observations in Figure 4.5A that both H1–5 and H4–5 can activate $USP7_{CD}$ to a similar extent.

4.4.5 Trans activation of $USP7_{CD}$ by HUBL domains persist when evaluated with a more native substrate

The substrate Lys^{48} di–ubiquitin was utilized to determine if the same activation trends persisted with a more native substrate. The data is summarized in Figure 4.7. The results show that $USP7_{FL}$ completely processed the Lys^{48} di–ubiquitin to mono–ubiquitin within the two hours, while no cleavage of Lys^{48} di–ubiquitin by $USP7_{CD}$ could be visualized. Mono–ubiquitin could be observed with only 2 μM of H1–5 *in trans* with $USP7_{CD}$, while 6 μM of H4–5 was required to observe mono–ubiquitin in the same amount of time. These results demonstrate that the cleavage of Lys^{48} di–ubiquitin increases with respect to the concentration of HUBL truncation present (Figure 4.7). Furthermore, these findings indicated that our experiments utilizing Ub–Rho110 as a substrate accurately reflect the mechanism of activation for USP7, and support the findings that H1–3 are required for the activation of $USP7_{CD}$.

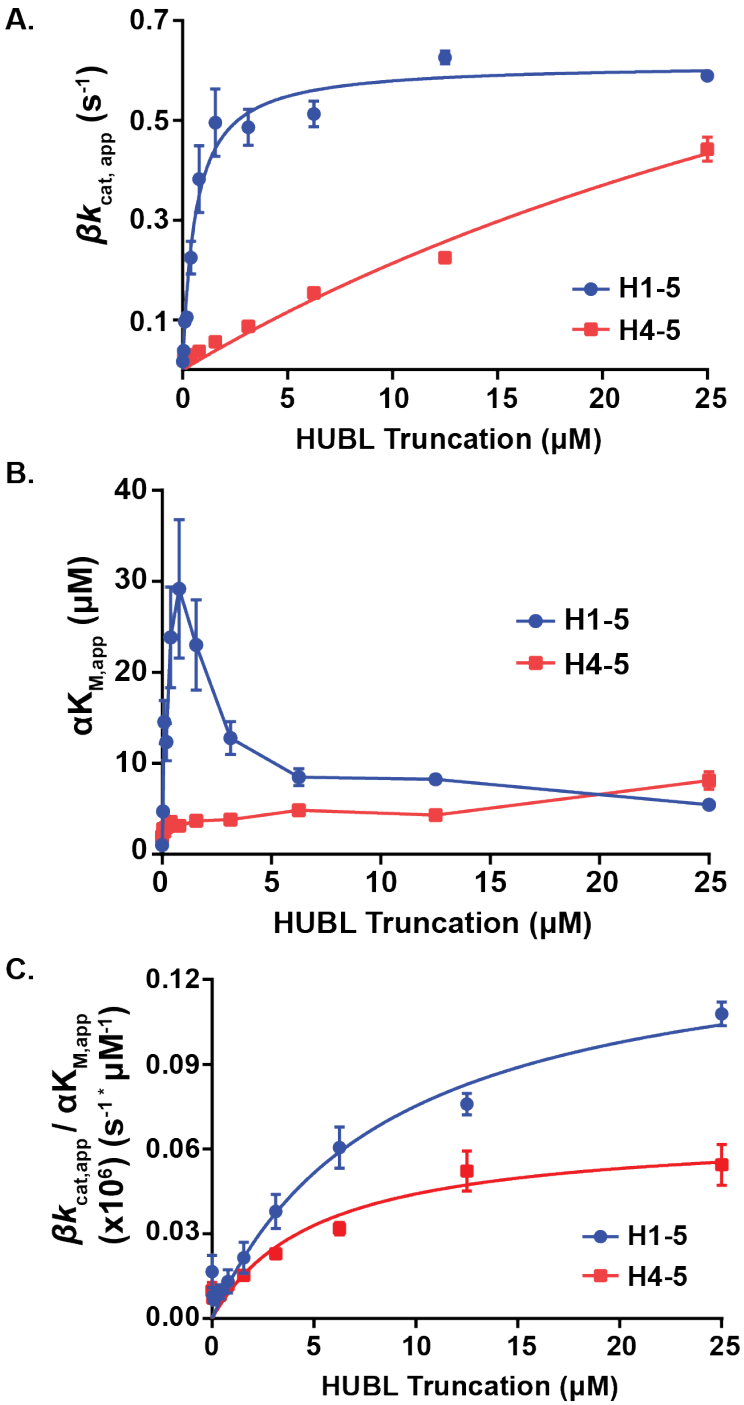


Fig. 4.6. **Modulation of Steady–State Kinetic Parameters of USP7$_{CD}$ as a function of HUBL Truncation Concentration.** The individual curves from Figure 4.4A and B were fit to the Michealis–Menten equation. **A.** From this fit, $\beta k_{cat,app}$ values in terms of increasing concentrations of H1-5 and H4–5 were determined. **B.** The variation in the $\alpha K_{m,app}$ values in terms of increasing H1–5 and H4–5 concentrations. **C.** Evaluation of the turnover of substrate by both systems in terms of HUBL truncation concentration as calculated from the fits of the individual curves from Figure 4.4A and B to the Michealis–Menten equation

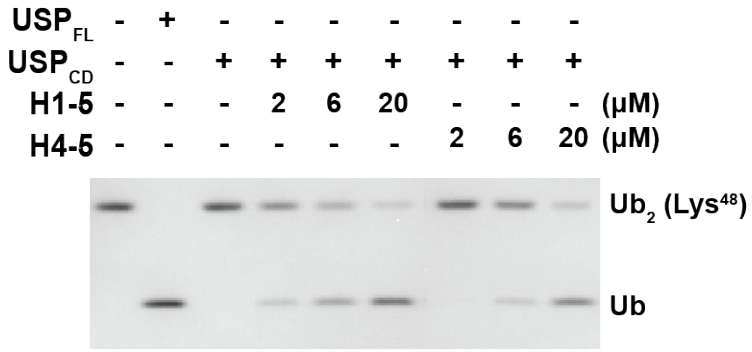


Fig. 4.7. **Cleavage of Di–Ubiquitin Substrate Lys^{48} by USP7$_{CD}$ with HUBL Truncations *in trans*.** Survey of the hydrolysis of 0.5 μg of di–ubiquitin substrate *Lys*^{48} by 50 nM of USP7$_{FL}$ (Lane 2), USP7$_{CD}$ (Lane 3), and USP7$_{CD}$ *in trans* with noted concentrations of either H1–5 (Lanes 4–6) or H4–5 (Lanes 7–9). Concentration of HUBL truncations varied as stated from 2 to 20 μM.

4.4.6 H1–3 can disrupt USP7$_{FL}$ and USP7$_{CD}$–H4–5 activity

To further evaluate the role of H1–3 in the activation of USP7$_{CD}$, HUBL truncations were added to USP7$_{FL}$ and Ub–Rho110. The percent inhibition was calculated from the measured activity (Equation (4.3)), and plotted in terms of the respective concentration of HUBL truncation (Figure 4.8A). When H4–5 was added to USP7$_{FL}$ it appeared to have no effect on the system, as is observed by the negative percent inhibition values. HUBL truncations H1–2 and H1–3 strongly interfered with the native activity of USP7$_{FL}$ and had IC$_{50}$ values of 45 $\pm$ 3 μM and 9 $\pm$ 1 μM respectively. H1–5 showed weaker interference with a calculated IC$_{50}$ value of 201 $\pm$ 40 μM. The larger IC$_{50}$ value associated with H1–5 may be a result of HUBL domains 4 and 5 being present within the truncation and being able to partially reactivate the system upon disruption.

We next attempted to disrupt USP7$_{CD}$–H4–5 with H1–3 and H4–5. Interestingly, we saw inhibition in both systems. Disruption by H1–3 resulted in an IC$_{50}$ value of 46 $\pm$ 3 μM which is slightly larger than that observed with USP7$_{FL}$ (Figure 4.8B).

Unlike $USP7_{FL}$, the activity of $USP7_{CD}$–H4–5 could be weakly disrupted by H4–5 with an IC_{50} value of 190 ± 35 μM (Figure 4.8C).

To evaluate if H1–3 interacts directly with H4–5 they were evaluated by size exclusion chromatography (SEC). First H1–3 was injected onto a 24 mL Sepharose 6 (GE Healthcare) chromatography column, and a peak eluted at 17.33 mL (Figure 4.9A, green). Second, H4–5 was injected onto the column and a peak eluted at 18.18 mL (Figure 4.9A, purple). Finally, both truncations were combined at a concentration of 65 μM in a 1:1 molar ratio to evaluate if they formed a complex (Figure 4.9B). Two peaks were apparent in the mixed injection (Figure 4.9B), first at 17.3 mL and second at 18.2 mL. Based on the results, no direct interaction appears to take place between H1–3 and H4–5. Together, these results further support a model where H1–3 is required to stabilize H4–5's interaction with $USP7_{CD}$, however it remains unclear how H1–3 dictates this stabilization. Furthermore, without H1–3 present, H4–5 is able to disrupt activity the native H4–5 in $USP7_{CD}$, an observation which was not observed with $USP7_{FL}$.

Table 4.3
Percent Interference of $USP7_{FL}$ by Various HUBL Truncations

HUBL Truncation	IC_{50} (μM)	Maximum % Inhibition
H1–5	201 ± 40	125 ± 14
H4–5	14 ± 14	-19 ± 5
H1–2	45 ± 3	100 ± 2
H1–3	9 ± 1	81 ± 3

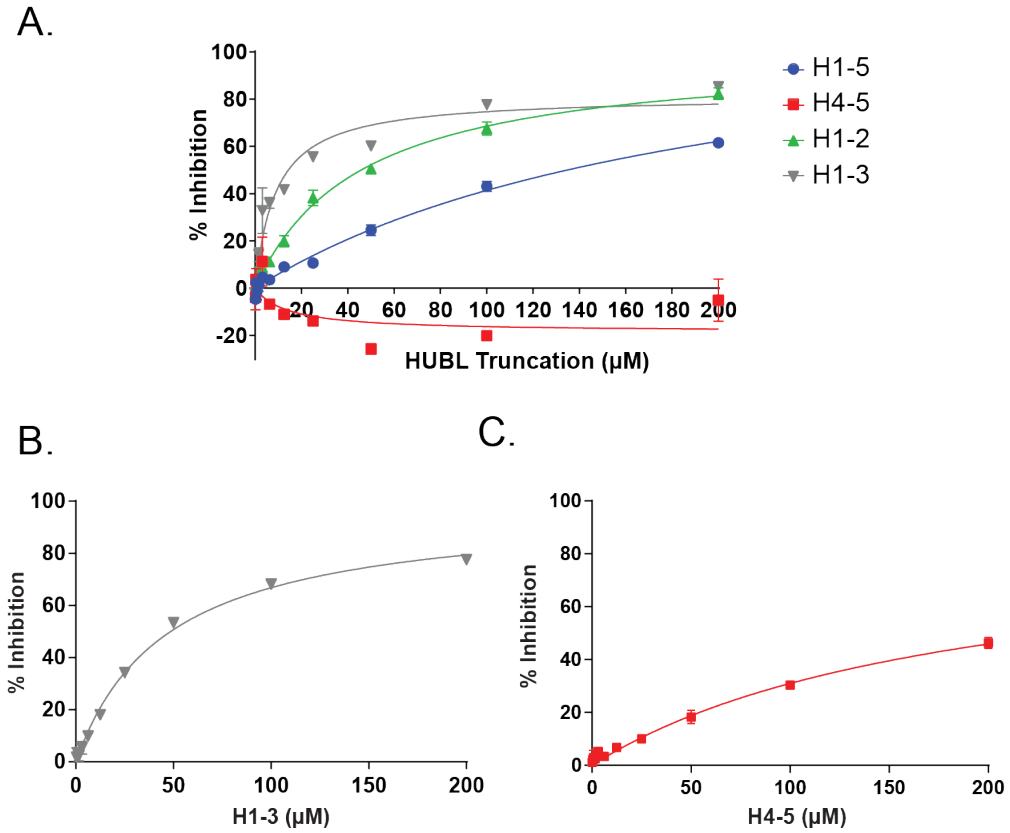


Fig. 4.8. **Interference of USP7$_{FL}$ Activity by HUBL Domains *in trans* A.** The percent inhibition of the activity of USP7$_{FL}$ as a result of HUBL truncations interference. Blue circles: H1–5. Red squares: H4–5. Green triangles: H1–2. Grey inverted triangles: H1–3. **B.** The percent inhibition of the activity of USP7$_{CD}$–H4–5 with H1–3. **C.** The percent inhibition of the activity of USP7$_{CD}$–H4–5 with H4–5. Error Bars, SD.

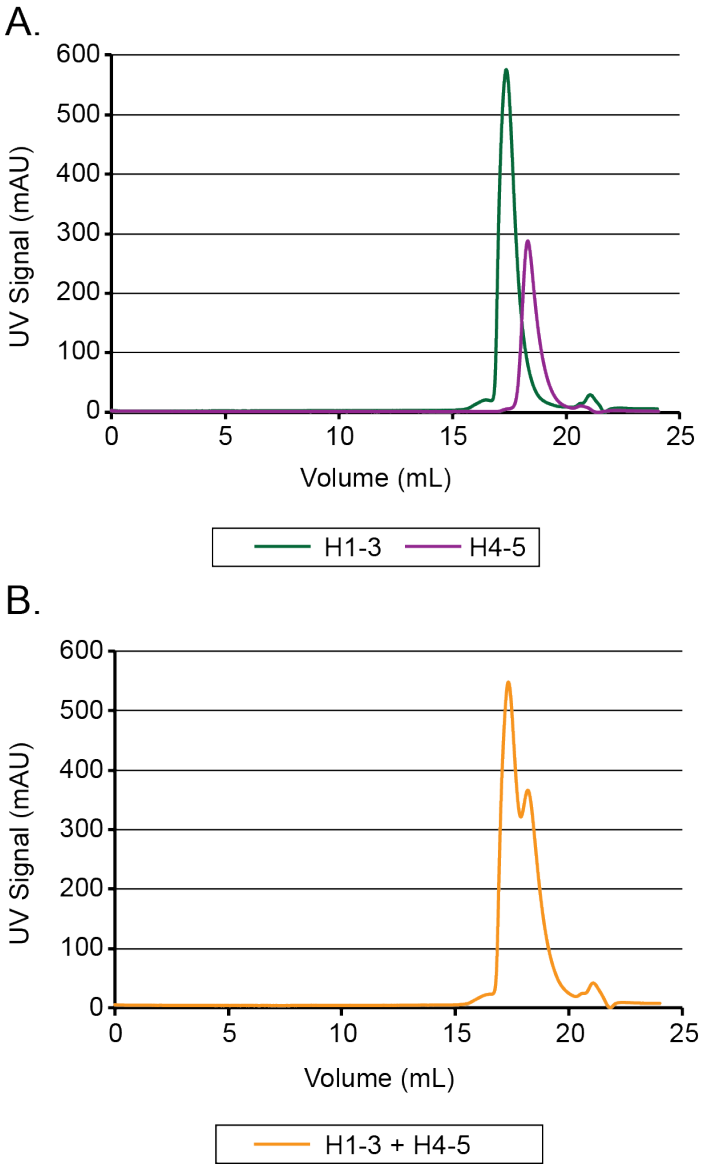


Fig. 4.9. **Evaluation of direct interaction between H1–3 and H4–5 by SEC.** H1–3 and H4–5 were run over a Sepharose 6 (GE Healthcare) size exclusion chromatography column. **A.** Chromatography curves of H1–3 and H4–5 run individually at a concentration of 65 μM. H1–3, green. H4–5, purple. **B.** H1–3 and H4–5 were combined at a concentration of 65 μM each in a 1:1 molar ratio. Resulting curve in orange.

4.5 Discussion

Several members of the USP family have been described to have at least one if not multiple UBL domains; however very little is known about the function of these UBL domains within the USPs. One of the most well studied functions of an UBL is within USP14. The UBL domain in USP14 appears to play an important role in the association of USP14 to the proteasome. This association results in the alleviation of the steric hindrance of the active site by the blocking loops [19]. In the case of USP7, previous work has shown that removal of the TRAF domain from $USP7_{FL}$ does not affect activity, but removal of the HUBL domains reduces activity substantially [34]. Structural and biochemical studies have described the importance of HUBL domains 4 and 5 in the activity of USP7 as well as elucidated a binding pocket for the C–terminus peptide of HUBL 5 within the catalytic domain of USP7 [20, 64]. While Faesen and coworkers did not find evidence for the importance of H1–3, the work presented here shows that H1–3 are required to regulate the activation by H4–5 to achieve the native form of intramolecular activation of USP7.

Kinetic analysis of $USP7_{CD}$ and $USP7_{FL}$ with Ub–Rho110 found that both have similar K_m values, but show great variation in their ability to process the substrate, as $USP7_{FL}$ is 30–fold more efficient than $USP7_{CD}$. This supports the findings that the HUBL domains are required for full intramolecular activation of USP7 [20, 34]. Investigation of the activity of $USP7_{CD}$ with H1–5 or H4–5 *in trans* exposed distinct differences in their ability to activate $USP7_{CD}$. These differences were also observed with Lys^{48} di–ubiquitin substrate, which more closely resembles the native substrate of USP7. The kinetic data was fit to the non–essential activation model and kinetic parameters α, β, K_A, K_m, k_{cat} were determined. The derived α–value for H4–5 activation of $USP7_{CD}$ was greater than one which indicates an increase in K_A and K_m in the presence of substrate. The opposite is true for H1–5 which has an α–value of less than one. Previous reports have described the interaction between H4–5 and $USP7_{CD}$ as easily disrupted by a single point mutation in the C–terminus loop of H4–5 [20,64].

The continuous increase observed in the $\alpha K_{A,app}$ value for H4–5 with increasing Ub–Rho110 concentration indicates that the presence of substrate further weakens the interaction of H4–5 with $USP7_{CD}$. With no substrate present, the K_A value for both truncations is identical. These differences in α–values strongly suggest that H1–3 has a role in stabilizing the interaction between H4–5 with $USP7_{CD}$, specifically in the presence of substrate.

The derived values of the kinetic parameter β suggests that the role of H1–3 exceeds that of just stabilizing the interaction between H4–5 and $USP7_{CD}$ in the presence of substrate. The β–value of H4–5 is 16–fold higher than that of H1–5. It appears that in the presence of H1–3, $USP7_{CD}$ cannot turnover substrate as efficiently as it can with just H4–5 alone. This suggests that H1–3 may attenuate USP7 activation by H4–5 in the presence of substrate. This hypothesis is supported by the observation that when introduced to $USP7_{FL}$, H1–3 can disrupt activity with an IC_{50} value of 9 $\pm$ 1 μM, the lowest IC_{50} value of the HUBL truncations tested. From these results, we propose that H1–3 plays the role of a rheostat in this system by regulating the activation levels $USP7_{FL}$ experiences by H4–5.

When comparing the enzymatic efficiencies of USP7 by the two truncations, it appears that the activation by H4–5 is not drastically different than that of H1–5. Only by careful kinetic characterization and interpretation of the α– and β–values associated with the non–essential activation model described above do we gain an understanding for how this system is controlled. When considering the α– and β–values of the H1–5 system separately from those values for the H4–5 system it becomes clear why there is only a 2–fold difference in enzymatic efficiency. The interaction between H4–5 and $USP7_{CD}$ is weakened in the presence of substrate, as is seen by the large α–value. However, the β–value for H4–5 is significantly larger than that with H1–5 activation and therefore this system is able to overcome the effects caused by the weakened interaction in the presence of substrate. The α– and β–values for H1–5 are as expected for an activation system, the $\alpha < 1$ strengthens the interaction of the $USP7_{CD}$–H1–5–Ubiquitin complex (Figure 4.2) while $\beta > 1$ increases the turnover of

the substrate. It is because of the large β value associated with the H4–5 activation of $USP7_{CD}$ that we do not see more drastic loss of enzymatic efficiency than we do. It is understandable that without such careful analysis, the importance of H1–3 may be overlooked in the activation of USP7.

Further evidence for the importance of H1–3 regulation was observed by differences in the interference caused by introducing H4–5 to $USP7_{CD}$–H4–5 *in trans* as compared to $USP7_{FL}$. Interestingly, without an internal H1–3 present H4–5 *in trans* was able to disrupt the activation provided by the tethered H4–5 in the $USP7_{CD}$–H4–5 construct. While the H4–5 *in trans* would still activate $USP7_{CD}$–H4–5, the activation level is significantly lessened without the tether and appears as inhibition within this system. When this same experiment was performed with $USP7_{FL}$, H4–5 was unable to disrupt activity and did not appear to have an effect on $USP7_{FL}$. Without H1–3 present in the $USP7_{CD}$–H4–5 construct, the regulatory role of H1–3 as defined in this paper was lost, and the burden of the large α–value associated with H4–5 *in trans* is eliminated. These results clearly support the proposed regulatory role of H1–3 has on H4–5 in the activation of USP7.

Currently there are two accepted orientations of the HUBL domains with respect to one another as determined by X–ray crystal structures [20,37,39]. The full structure of the HUBL domains shows the domains in an elongaged 2–1–2 fashion while the structures of HUBL domains 1–3 (with or without the catalytic domain) show a compact structure due to a hinge between HUBL domains 2 and 3. From these structures it was hypothesized that the HUBL domains fold back toward the catalytic domain for proper alignment for catalysis. The kinetic data discussed here not only supports this hypothesis, but further suggests that the compact state of the HUBL domains is essential to allow for intramolecular interactions between the domains which dictates USP7 activation, a finding that could not be elucidated from the available structures alone.

The X–ray structure of the catalytic domain with HUBL domains 1–3 displays the importance of the α–helix of the catalytic domain (residues 535–560) for the proper positioning of the HUBL domains with respect to the catalytic domain [63].

Utilizing this structure, as well as biochemical studies, Kim and coworkers developed a model for the regulatory function of the tether in regards to its rigidity, length, and charge localization to spatially align the HUBL domains to the catalytic domain. They propose that the α–helix tether positions the HUBL domains at the proper distance from the catalytic domain so that all five HUBL domains are required to fold back comfortably, without steric hindrance, toward the catalytic domain to properly align the C–terminal peptide of HUBL 5 for activation [63]. Support for the proper spatial alignment of the C–terminal peptide of HUBL 5 was expressed with kinetic studies performed Rougé and coworkers where they defined the appropriate linker length for activation of $USP7_{CD}$ by just the C–terminal peptide. When they created a construct of $USP7_{CD}$ tethered to the C–terminal peptide of HUBL 5 with a 10 or 20 amino acid linker, the k_{cat} values measured were within 2–fold of $USP7_{FL}$ [64]. If they shortened the linker, the k_{cat} values increased. These results, however, do not explain the similarity of the k_{cat} values determined with our $USP7_{CD}$–H4–5 construct in comparison to the k_{cat} values determined with $USP7_{FL}$.

Figure 4.10A summarizes what is known of the different states USP7 can assume. Under basal conditions in the cell, the HUBL domains exist in equilibrium between a not–engaged (top) and engaged (bottom) state with the catalytic domain. When the HUBL domains assume the engaged state they interact with the catalytic domain causing USP7 to enter an activated state which results in an increase in k_{cat}. With the available X–ray crystal structures, we have snapshots of the apo enzyme without HUBL domain engagement (PDB: 4M5W, 1NB8) [15,16], apo enzyme with potential HUBL domain engagement (Kim et al, and PDB 5J7T) [63, 64], and the ubiquitin bound catalytic domain without HUBL engagement (PDB: 1NBF) [16]. Furthermore, we have now have an idea of what HUBL engagement with the ubiquitin bound catalytic domain may look like with the C–terminal peptide of HUBL 5 bound in an

activation cleft of $USP7_{CD}$ [64]. The catalytic triad is misaligned in all structures lacking ubiquitin and is aligned in all structures containing ubiquitin. Until recently, how the HUBL domains 4–5 activate USP7 was unclear. It had previously been proposed that this activation was a result of an interaction of the C-terminal peptide of the HUBL domain 5 with the switching loop of the catalytic domain [20]. However, Rougé and coworkers recently published two X–ray crystal structures suggesting that upon ubiquitin binding to the catalytic domain, an activation cleft opens [64]. This within this activation cleft, the C–terminal peptide of HUBL 5 was bound. This suggests that the activation observed with H4–5 is a result of stabilization of the ubiquitin bound construct and is indicative of an induced fit model [64].

GMPS is an allosteric activator of USP7. When the cell experiences genotoxic stress, GMPS is translocated to the nucleus and directly interacts with USP7 by binding to the acidic pocket of HUBL domain 2 [37,39]. Through this interaction, GMPS is able to allosterically activate USP7 by further increasing $\beta k_{cat,app}$ to over 5–fold what is observed by the intramolecular activation of the HUBL domains alone [20]. Based on the research presented here, this finding supports our hypothesis that H1–3 regulates USP7 catalytic activity by maintaining a lower β–value than what is seen with just H4–5 with $USP7_{CD}$ *in trans*. The binding of GMPS to H1–3 may alleviate the regulation imposed by H1–3, allowing for an increase in the β–value and higher activity of USP7. The physiological ramifications of this activation in the cells response to stress supports this hypothesis. More p53 would be stabilized in a less amount of time leading to a stronger and faster downstream response. Further analysis will need to be done to determine the exact mechanism by which H1–3 regulates H4–5 in the intramolecular activation of USP7.

By combining the kinetic results presented here with the available X–ray crystal structures of the HUBL domain orientations, we propose the following rheostat model for the intramolecular activation of USP7 by its HUBL domains (Figure 1.4B) [20, 39, 63]. In the not-engaged state (1), the HUBL domains do not associate with the catalytic domain, and, as such, the activity is 30–fold less than the engaged

state. When USP7 assumes the engaged state (2) the HUBL domains 1–3 act as a rheostat to dial up the activity by stabilizing and regulating the interaction of HUBL domains 4–5 with the catalytic domain. While H4–5 have the potential to activate USP7 further, H1–3 act as a regulator, and hold this activity in check. When the cell experiences genotoxic stress, the allosteric activator, GMPS, translocates to the nucleus and activates USP7 by interacting with HUBL domains 1–3 (3) to alleviate the regulation of H1–3 on H4–5 while maintaining stabilization with the catalytic domain to increase activity levels over 5–fold than without GMPS present. This model further describes the importance of H1–3 in the intramolecular activation of USP7 and catalysis.

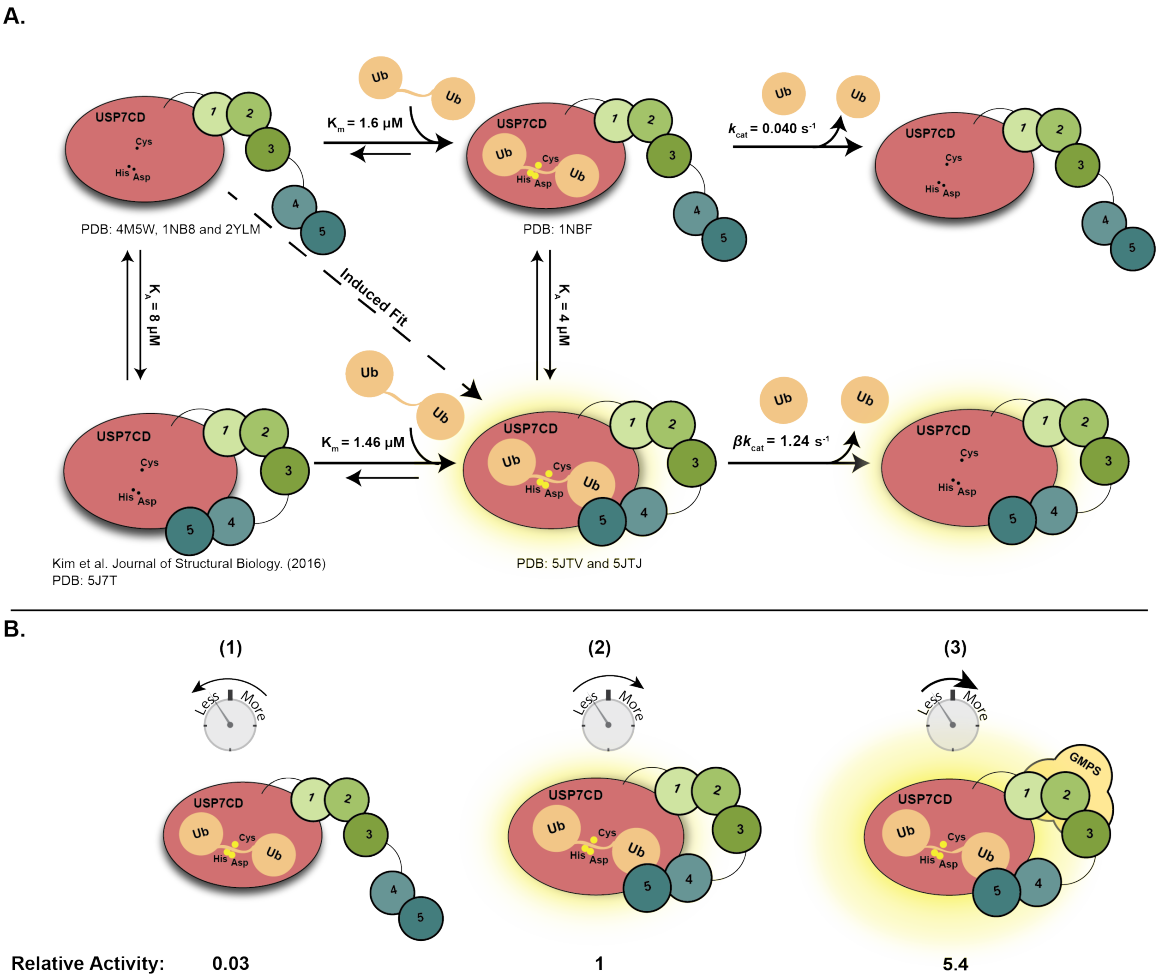


Fig. 4.10. **Tethered–Rheostat Model for the Intramolecular Activation of USP7.** HUBL domains 1–3 act as a rheostat to dial the activity USP7 experiences to more or less depending on the state of the cell. **A.** $USP7_{CD}$ (pink) exists in an equilibrium between not–engaged (Top Row) and engaged (Bottom Row) HUBL domains. In the not–engaged state, the HUBL domains do not influence activity and the k_{cat} value is minimal. In the engaged state, H1–3 regulate the interaction of H4–5 with the catalytic domain to regulate the activity of USP7 through manipulation of the kinetic parameters to increase the k_{cat} value by the β–value. **B.** USP7 exists in three states, (1) HUBL domains not engaged, (2) HUBL domains engaged, and (3) USP7 in complex with GMPS. When the HUBL domains are not engaged (1) with the catalytic domain the activity is significantly lower than when they engage with the catalytic domain (2) as shown by the relative activity. Based on the data presented here, we propose that H1–3 regulate and stabilize the interaction of H4–5 with the catalytic domain. However, H1–3 also regulates the amount of activity USP17 experiences. When HUBL domain 2 binds GMPS (3) the interaction alleviates the negative regulation imposed on H4–5 by H1–3 and allows for an elevated activity state. In this way, H1–3 act as a rheostat to regulate USP7's activity.

CHAPTER 5. HIGH-THROUGHPUT SCREENING AND HIT COMPOUND EVALUATION FOR USP17

5.1 Introduction

As described in Chapter 2, USP17 regulates cell–cycle progression by the deubiquitination of the Ras and CDC25A pathways. When the balance of either of these pathways is perturbed, cell proliferation and hence cancer can result. The persistent overexpression of USP17 has been attributed to promotion of prostate, breast, hematopoietic, and non–small cell lung cancers [27, 29, 46]. Previous work has validated USP17 as a target for the development of anti–cancer theraputics since knocking down overexpressed USP17 by shRNA leads to a reduction in the size of breast cancer xenographs [29]. The expression and purification methodology for obtaining high yields of USP17 described in Chapter 2 allowed for the development of high–throughput screening (HTS) assays for the identification of small molecule inhibitors towards USP17 which is described herein.

This chapter describes the development and optimization of a high–throughput screening assay for USP17. The assay was utilized to screen 25,000 compounds from the ChemBridge library composed of drug–like scaffolds. The identified scaffolds were used for development of structure activity relationship described within.

5.2 Experimental Procedures

5.2.1 High-throughput screening (HTS) of a diverse 20K library

HTS Assay Design

The purchased library used to screen USP17 contained 25,000 compounds containing drug–like scaffolds that were cherry–picked from the entire ChemBridge Corporation collection of compound libraries. The compounds are stored at a concentration of 10 mM in 100% dimethyl sulfoxide (DMSO) at -20 °C. The kinetic assay described in Section 2.3.12 was redesigned as an end–point assay, and scaled down to a final volume of 5 μL for use in 1536–well plates. The screen was automated with the use of a ThermoScientific Multidrop Combi reagent dispenser (for addition of the enzyme and substrate) and a BiomekFX liquid handling system (for addition of the compounds). To perform the screen, 4 μL of 1.25 nM USP17 (a final concentration of 1 nM) in Screening Buffer (50 mM CHES pH 9.5, 5 mM DTT, 0.1 mg/mL BSA, and 0.01 % Trition X–100) was dispensed into a Nunc 1536–well High Base, black polystyrene microplate. The compound was added in a volume of 20 nL dispensed with a VP Scientific 384–well pin–tool attached to a BiomekFX liquid handling system (a final concentration of 10 μM). The reaction was initiated with 1 μL of 1μM Ub–Rho110 suspended in Screening Substrate Buffer (50 mM CHES pH 9.5, 1 mM EDTA, 100 mM NaCl, 0.05% CHAPS) to a final concentration of 200 nM, and allowed to incubated for 20 minutes. After incubation, 1 μL of 0.5 M Acetic Acid (a final concentration of 83 μM) was added to quench the reaction. The end–point fluorescence was then read using a Beckman–Coulter DTX880 plate reader at an excitation λ of 485 nm and emission λ of 535 nm. The screen was performed in duplicate using separate 1536 well plates at room temperature. The screen set up is depicted in Figure 5.1. Each plate contained two columns (64 wells) of USP17 without compound present and two columns of substrate alone for determination of *Z'*–factors.

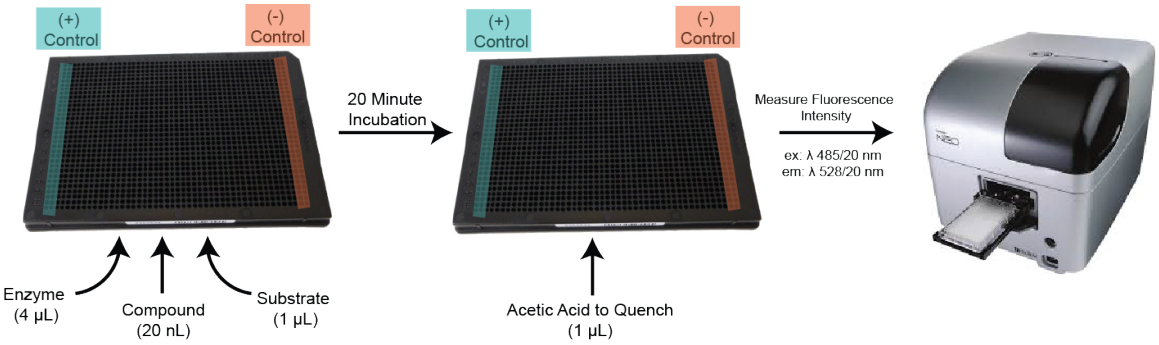


Fig. 5.1. **HTS workflow using 1536–well microplate format.** For the USP17 screen, each plate contained 128 control wells, as depicted with blue and red shading, and 1,408 sample wells. The assay was performed in 3 main steps. (1) the enzyme, compound, and substrate were added to the plate. (2) the plate is incubated for 20 minutes and quenched with acetic acid. (3) the fluorescence intensity was measured.

Z–factor and *Z'*–factor Determination

The Z–factor is a simple and dimensionless parameter for evaluating the quality of an HTS screen. It is calculated using Equation (5.1) to quantify the ratio between the separation band (the signal between the sample reading and the control reading) and the dynamic range (the absolute value of difference between the sample signal, μ_s, and the control signal, μ_c means) as depicted in Figure 5.2. The separation band is the sum of three standard deviations of the sample ($3\sigma_s$) and the control ($3\sigma_c$). A Z–factor value of ≥ 0.5 indicates a large separation between the sample and separation bands indicating the assay is of high quality.

$$\mathrm{Z} = 1 - \frac{(3\sigma_\mathrm{s} + 3\sigma_\mathrm{c})}{|\mu_\mathrm{s} - \mu_\mathrm{c}|} \tag{5.1}$$

The values corresponding to sample are defined as the signal from the enzyme in the presence of a compound and the values defined as the control are calculated from the signal of the enzyme without inhibitors present. The Z–factor is typically calculated using a well–characterized inhibitor of the enzyme and the uninhibited

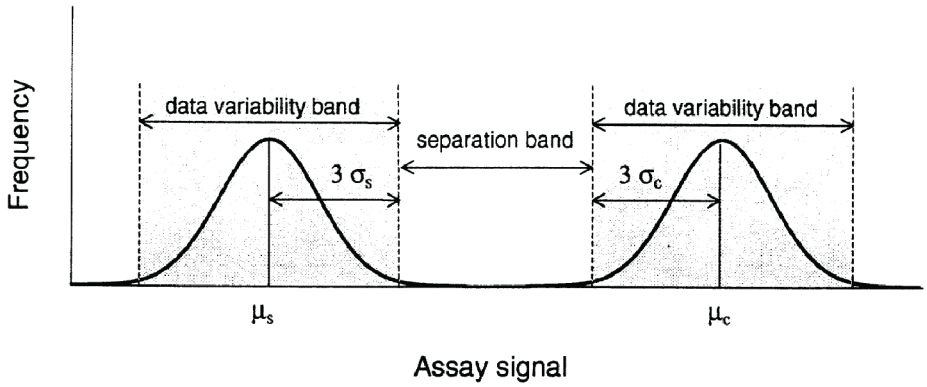


Fig. 5.2. **Illustration of *Z*–factor determination.** Depiction of the data variability band measurement for the control and the sample as well as the 3σ and μ values for the control and the sample. The separation band measurement is also depicted. These measurements are used to determine the *Z*– and *Z'*–factors for HTS. Figure adapted from [66].

control, however in the case of USP17 no well–characterized inhibitor is known [66]. Therefore the *Z'*–factor was calculated using Equation (5.2) using only the control data of the signal of the reaction with enzyme ($3\sigma_{c+}$ and μ_{c+}) and without enzyme ($3\sigma_{c-}$ and μ_{c-}) present [66].

$$Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|} \tag{5.2}$$

5.2.2 Hit–2–Lead Follow–up Analysis

Preparation of Hit–2–Lead compounds

Three promising scaffolds were identified from the screening (Figure 5.3) and 13 compounds containing these scaffolds were ordered as their dried powder from ChemBridge Hit–2–Lead Online Chemical Store for further analysis (Figure 5.4). Each compound was dissolved in 100% DMSO to a stock concentration of 40 mM. Com-

pounds **8** and **10** were not soluble at 40 mM and therefore were diluted to 10 mM in DMSO.

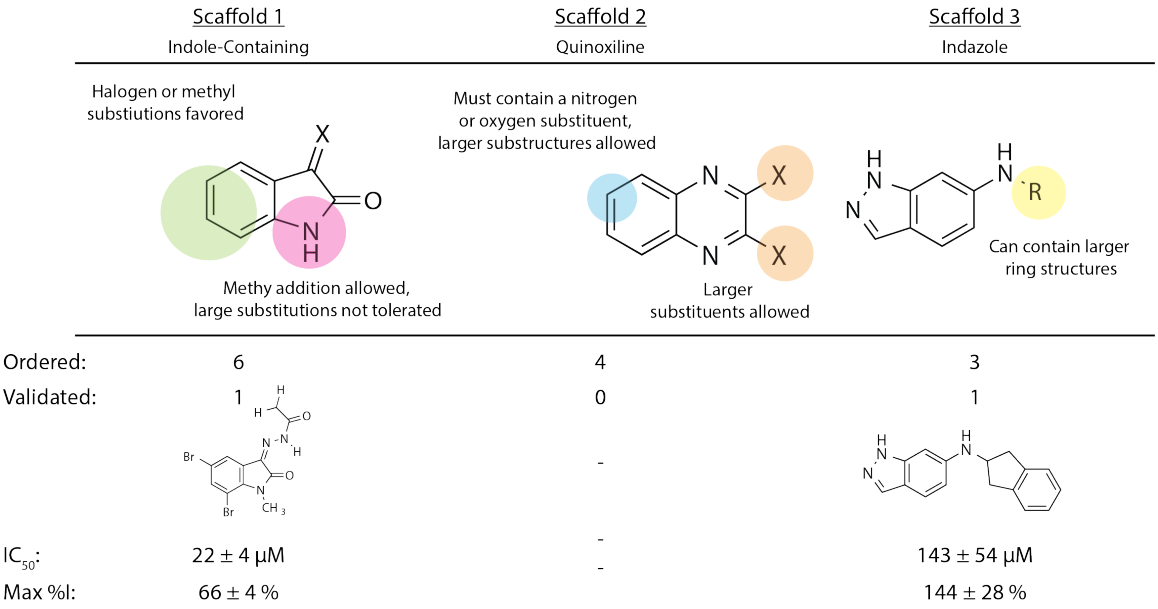


Fig. 5.3. **USP17 HTS Summary.** Three compound scaffolds were identified from the original HTS, and the follow–up screen with the Diverset ChemBridge library that showed promising structure activity relationships, as described. The purchased value is the number of compounds from each scaffold that were purchased from Hit–2–Lead. The validated value is the number of compounds that IC_{50} values could be measured for. The validated hits are shown along with their corresponding IC_{50} value.

% Inhibition of Hit–2–Lead compounds

The % inhibition of each compound was determined prior to determining the IC_{50} values to conserve substrate. The final concentration of substrate was held constant at 100 nM and the final enzyme concentration was held at a final concentration of 1 nM, which is identical to that of the original screen. From the 40 mM stock, each compound was diluted to a working concentration of 5 mM in 100% DMSO. The assay was performed as follows: 1 μL of the working stock of compound was added to a Costar 96 half–volume black plate, and mixed with 15 μL USP17. Plates were

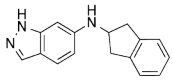
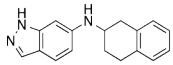
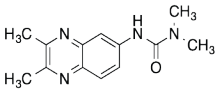
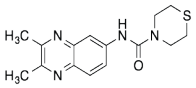
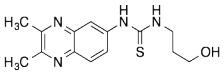
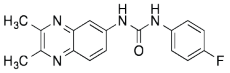
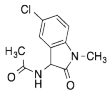
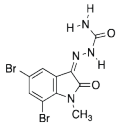
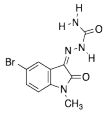
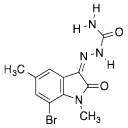
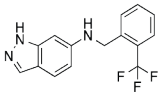
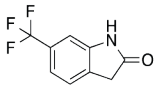
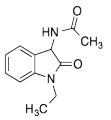
Compound Name	Structure	% Inhibition	IC_{50} (μM)
Indazole-1		24 ± 7 %	22 ± 4
Indazole-2		14 ± 7 %	N/A
Quinoxilline-3		23 ± 8 %	N/A
Quinoxilline-4		4 ± 5 %	
Quinoxilline-5		N/A	
Quinoxilline-6		N/A	N/A
Indole-7		N/A	
Indole-8		97 ± 2 %	66 ± 4
Indole-9		N/A	N/A
Indole-10		9 ± 5 %	N/A
Indazole-11		N/A	N/A
Indole-12		N/A	
Indole-13		N/A	

Fig. 5.4. **Hit–2–Lead Compound Summary.** Compounds are labeled numerically with identifying scaffold. The % inhibition was tested for all 13 compounds at 100 μM. IC_{50} measurements were attempted for 8 of the compounds, of which 2 were successful (green boxes).

allowed to incubate for five minutes at room temperature. To initiate the reaction, 14 μL of Ub–Rho110 was added. Each assay was measured in triplicate. A negative control of Ub–Rho110 alone was performed to evaluate the background fluoresence. All reactions contained a final concentration of 2% DMSO. The enzymatic activity was measured as described in Section 2.3.12, and the % inhibition was calculated as follows:

$$\% \text{ Inhibition} = \left(1 - \frac{\text{Rate}_{\text{inhibited}} - \text{Rate}_{\text{neg}}}{\text{Rate}_{\text{uninhibited}} - \text{Rate}_{\text{neg}}}\right) \times 100 \tag{5.3}$$

In Equation (5.3), $Rate_{inhibited}$ is the measured initial slope of the progress curve of USP17 in the presence of inhibitor, while $Rate_{uninhibited}$ is the measured initial slope of the progress curve of USP17 in the absence of inhibitor. $Rate_{neg}$ is the slope of the background fluorescence with respect to time in the absence of enzyme.

Determination of IC_{50} values

Only 8 of the 13 Hit–2–Lead compounds showed measurable inhibition of USP17 at 100 μM, therefore, these compounds were further evaluated to determine their IC_{50} values. The IC_{50} value refers to the concentration of compound required to reach half–maximal inhibition of the enzyme. To determine the IC_{50} values, eleven serial dilutions of each compound of interest were prepared from the 40 mM stock to working stock concentrations of 10 mM to 20 μM in 100% DMSO. 1 μL of working stocks of compound were added to a Costar 96 half–volume black plate and 15 μL of USP17 (a final concentration of 1 nM) was added to each well. The plate was mixed and incubated at room temperature for five minutes. To initiate the reaction, 14 μL of Ub–Rho110 (a final concentration of 200 nM) was added to each well. The initial slope of the progress curves were measured, and the % inhibition was calculated with Equation (5.3). The resulting % inhibition was plotted as a function of inhibitor concentration. The IC_{50} was determined from the fit of each curve to Equation (5.4)

using non–linear regression analysis with the Enzyme Kinetics Module in the program SigmaPlot (v13:Systat Software Inc.).

$$\%\ \text{Inhibition} = \frac{\text{Max }\%\text{Inhibition} * [\ \text{Inhibitor}]}{[\text{Inhibitor}] + \text{IC}_{50}} \tag{5.4}$$

The hill equation was used to determine if cooperativity plays a role in the inhibition of USP17 by compound **8**. The Hill equation is as follows:

$$V = \frac{\text{V}_{\text{max}} * [\text{S}]^{\text{n}}}{\text{K}_{\text{m}}{}^{\text{n}} + [\text{S}]^{\text{n}}} \tag{5.5}$$

where V is the reaction velocity and V_{max} is the maximum reaction velocity. [S] is the substrate (Ub–Rho110) concentration and K_m is the Michealis–Menten coefficient. n is the hill coefficient. A n value of one indicates no cooperativity, > 1 indicates positive cooperativity, and < 1 indicates negative cooperativity.

5.3 Results

5.3.1 High–throughput screening optimization

The kinetic assay previously described for USP17 (Chapter 2, Section 2.3.12) was adapted to develop a 5 μL end–point assay for use with 1536–well plates. This optimization reduces the amount of enzyme and substrate required, as well as the length of time required to complete the entire screen. Variables that were optimized were enzyme and substrate concentrations, length of reaction time, and *Z'*–factors. Since the screening was performed in collaboration with the Biomolecular Screening and Drug Discovery Facility in the Purdue Discovery Park, all optimizations were done with their equipment. The goal of optimizing the enzyme and substrate concentration was two–fold. First, by reducing the amount of enzyme and substrate required the cost of the total screen would be significantly reduced. Second, manipulation of the enzyme and substrate concentrations varied the length of time, the reaction could proceed prior to quenching and reading the resulting fluorescence. Optimization of

this time period allowed for numerous plates to be run simultaneously, decreasing the time required to complete the entire screen. The determined optimal concentrations were 1 nM USP17 and 100 nM Ub–Rho110. This assay composition allowed the reaction to run for 20 minutes prior to adding a final concentration of 80 μM Acetic Acid to quench the reaction.

Prior to performing the entire screen, a single plate was tested in triplicate, and a *Z'*–factor of 0.8 was determined. After the first day of screening, two plates were evaluated and a *Z'*–factor of 0.5–0.6 was calculated. While still within the acceptable range, this drop in *Z'*–factor was cause for concern, and led to the reevaluation of how the screen was designed. Many variables were reanalyzed to attempt to increase the *Z'*–factor including making fresh reagents daily, sterile filtering all reagents, timing of the reaction, and washing the dispenser lines between plates to ensure accurate dispensing. Unfortunately, none of these adjustments resulted in an increase in the *Z'*–factor, therefore USP17 was screened as described above. As a *Z'*–factor of 0.5–0.6 is still within the acceptable range which indicates a strong assay, the entire library (a total of 41 plates) was screened over three days time. The resulting *Z'*–factors varied between 0.53 and 0.60 throughout the screen.

5.3.2 High–throughput screening evaluation and development of structure activity relationships

Figure 5.5 displays the percent inhibition mean and standard deviation calculated for the entire screen. Of the 25,000 compounds, 48 compounds consistently had percent inhibition values $\geq$ -1000% which may indicate that those compounds may activate the enzyme, but it is more likely that the compounds are auto–fluorescent. The percent inhibition cut–off was set to 38% to maintain a hit rate of less than 0.1%. With this inhibition threshold, 15 compounds were identified as potential hits. These identified hits fell into four classes: indole–containing (Figure 5.3, scaffold 1), indole like, quinoxiline substructure (Figure 5.3, scaffold 2), and nitroso containing

structures. The indole containing class had seven compounds, the indole like class had two compounds, the quinoxiline substructure class only had one compound, and the nitroso containing class had five compounds. Structure Activity Relationships (SARs) were determined for the indole–containing and the indole like compounds. To expand on the indole-containing and indole like SAR, as well as develop an SAR for the quinoxiline substructure, five plates of the ChemBridge Diverset library known to contain compounds of similar scaffolds based on substructure analysis, were screened in triplicate identically to the original screen. The nitroso containing compounds were not followed up with at this time due to toxicity commonly associated with these type of compounds. The results of the secondary screen supported the primary screen, and expanded the SAR of the indole–containing and quinoxiline substructure compounds. From this secondary screen an indazole–containing scaffold was identified, and was pursued further by ordering a series of compounds from the ChemBridge Hit–2–Lead library. The three promising scaffolds are displayed in Figure 5.3.

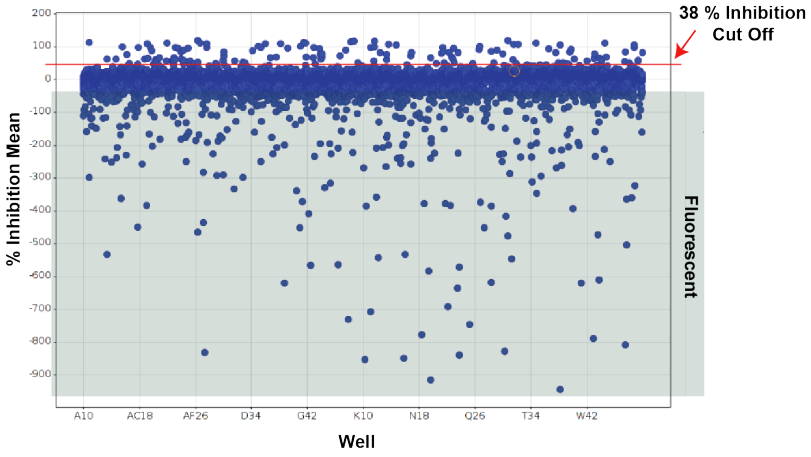


Fig. 5.5. **HTS results for primary screen for USP17 inhibition.** Distribution of the % inhibition data from the USP17 HTS screen. Red line represents the 38 % cut–off for inhibition. Grey shaded region represents compounds that showed increased fluorescence in the screen.

5.3.3 Follow–up Hit–2–Lead compounds for further development of SAR

Virtual screening of the identified hits by Dr. Laura Kingsley aided in the selection of compounds to be purchased to further develop the SAR of the indole–containing (scaffold 1), quinoxiline substructure (scaffold 2), and the indazole (scaffold 3) compounds. A total of 13 compounds were purchased ChemBridge Hit–2–Lead library (Figure 5.3), six containing scaffold 1, four containing scaffold 2, and three containing scaffold 3. The compounds were dissolved in 100% DMSO to 40 mM and the percent inhibition at 100 μM was determined for each compound (Figure 5.6). Compounds that showed measurable inhibition (**1**, **2**, **3**, **6**, **8**, **9**, **10**, **11**) were evaluated to determine IC_{50} values. Of these eight compounds, only two (compounds **1** and **8**) had measurable IC_{50} values (i.e. a dose–response curve was apparent and could be fit to Equation (5.4), Figure 5.7 and Figure 5.8). Compound **1** was part of the scaffold 3 group, and had an IC_{50} value of 22 $\pm$ 4 μM with a maximum % inhibition value of 66 $\pm$ 4 %. Compound **8** was part of the scaffold 1 group, and had an IC_{50} value of 144 $\pm$ 54 μM with a maximum % inhibition value of 144 $\pm$ 28 %. The IC_{50} curve for compound **8** appears to be sigmoidal in nature, and was refit using the Enzyme Kinetics Module of SigmaPlot (v13: Systat Software Inc.) to the Hill form of the Michealis–Menten equation (Equation (5.5)), and is displayed in Figure 5.8B. The determined Hill coefficient was n = 1.4 $\pm$ 0.4. With the large error associated with the hill coefficient, cooperativity of compound **8** cannot be determined by this analysis.

5.4 Discussion

The ChemBridge cherry–pick library was the first library to be screened because of the high number of deubiquitinating enzymes in the Mesecar Lab that had previously been screened against this library, including USP7 (Chapter 6). These previous screens provided a wealth of knowledge regarding promiscuous inhibitors that would save time and money in the identification of USP17 specific inhibitors. After ruling

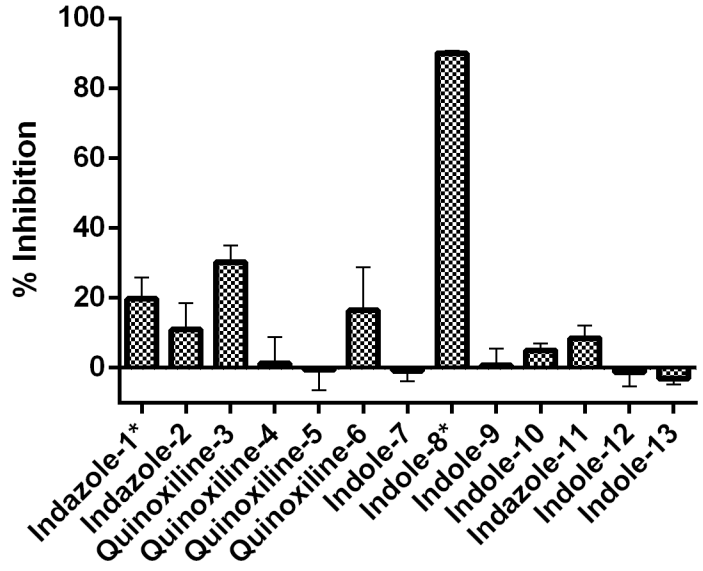


Fig. 5.6. **% Inhibition of Hit–2–Lead compounds.** Bar graph of the % inhibition of the 13 Hit–2–Lead compounds at 100 μM with 1 nM USP17 and 200 nM Ub–Rho110. Compounds are labeled numerically with identifying scaffold. Asterisks denote compounds with determined IC_{50} values. Error bars, SD.

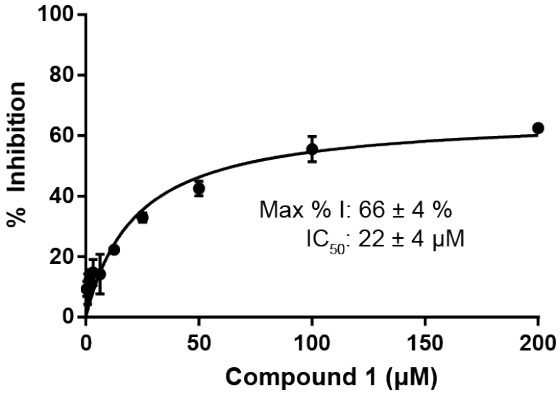


Fig. 5.7. **Compound 1 IC50 curves for USP17.** IC_{50} values were determined from dose–dependent % inhibition of USP17 after five minute incubation with compound **1** at room temperature. Error bars, SD.

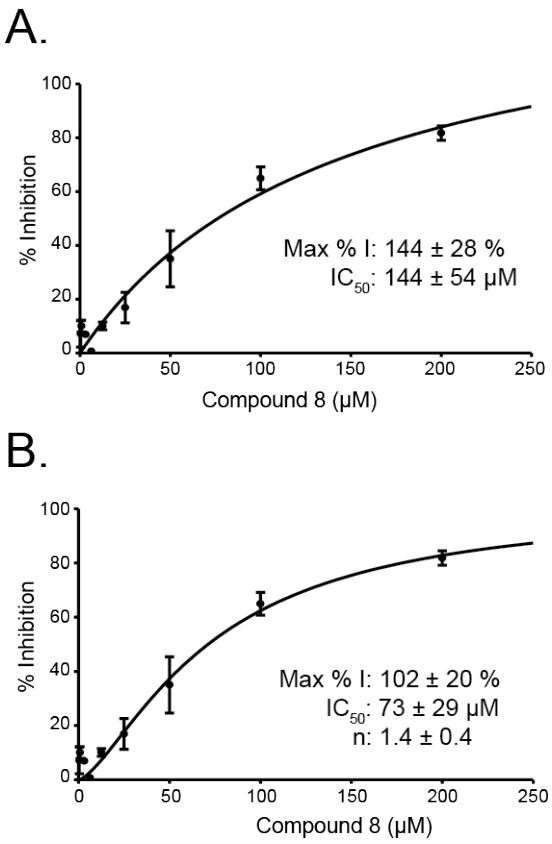


Fig. 5.8. **Compound 8** IC_{50} **curves for USP17. IC50** values were determined from dose–dependent % inhibition of USP17 after five minute incubation with compound **8** at room temperature. **A.** IC_{50} curve fit to Equation (5.4). **B.** Cooperativity analysis by fitting the IC_{50} curve to Equation (5.5) Error bars, SD.

out promiscuous compounds and compounds containing reactive nitroso groups, 15 hits remained. The hits could be classified into three scaffolds: indole-like, quinoxiline substructure, and indazole containing compounds (Figure 5.3). Thirteen compounds were ordered from the ChemBridge Hit–2–Lead library to further develop SAR (Figure 5.4), however only two of these compounds had measurable IC_{50} values towards USP17 (Figure 5.7 and Figure 5.8).

Compound **1** is of the indazole scaffold group and had a reasonable IC_{50} value of 22 $\pm$ 4 μM. The maximum % inhibition value of compound **1** was 66 $\pm$ 4 %, suggesting

that compound **1** may be a partial inhibitor. These values suggest that compound **1** would be promising to optimize further to increase its potency, and further develop the SAR with USP17.

Compound **8** is of the indole scaffold group, and had a weaker IC_{50} value compared to compound **1** at 143 $\pm$ 53 μM, but it demonstrated a higher maximum % inhibition than compound **1** at 144 $\pm$ 28 %. However, this higher maximum % value may be to due to the large amount of error seen associated with the IC_{50} curve. Furthermore, the IC_{50} curve of compound **8** appears to be sigmodial in nature which may indicated cooperativity. However, when the IC_{50} curve of compound **8** was fit the Equation (5.5) the Hill coefficent was 1.4 $\pm$ 0.4, so cooperativity within the system could not be determined.

It is unfortunate that we were only able to validate two compounds from this initial round of screening. Furthermore, concerns regarding the screen itself and the lower *Z'*–factors discussed in Section 5.3.2 suggest more work needs to be done. Taking this in mind, further optimization of the screen design is necessary to improve the *Z'*–factors. Furthermore, this screen took place prior to the pH evaluation of USP17 described in Chapter 2. The results from that experiment suggest that USP17 is most active between pH 7 and 9, therefore the Screening Buffers utilized here should be eliminated and the typical Assay Buffer described in Section 2.3.12 should be utilized for future screens. Finally, it would be ideal to branch out and screen USP17 against different libraries available within the Biomolecular Screening and Drug Discovery facility.

CHAPTER 6. IDENTIFICATION, OPTIMIZATION, AND MECHANISTIC EVALUATION OF A NOVEL ALLOSTERIC INHIBITOR OF USP7

6.1 Introduction

Deubiquitinating enzymes (DUBs) are involved in many different signaling pathways within the cell [13]. Tight regulation of DUB expression is required for normal cell progression as a large number of the 100+ DUB family members regulate cancer–associated pathways [13]. A common characteristic in many cancers is the upregulation of specific DUBs, and the subsequent downstream perturbation of their corresponding signaling pathways. Therefore, DUBs whose overexpression has been well–characterized in cancer may be are ideal targets for the development of small molecule inhibitors for cancer treatment [13].

Targeting an individual member of the DUB protease family was originally assumed to not be feasible due to the well–conserved cysteine and histidine boxes and fold of the catalytic domain described in Chapter 1. This is a similar thought process to what previously slowed drug discovery for the kinase enzyme family as well. However, the development and approval of the tyrosine kinase inhibitor, imatinib, for the treatment of BCR–ABL associated cancers in 2001 changed this way of thinking [67]. Since the approval of imatinib, 28 more specific kinase inhibitors have been approved by the FDA, the bulk of which were approved between 2012 and 2015 [67]. With such a high level of success targeting specific members of the kinase family, targeting specific members of the USP family may be more feasible than originally thought.

USP7 is responsible for maintaining the careful balance between MDM2 and p53 levels in the regulation of cell survival, as described in Chapter 4. Disruption of the p53 signaling pathway is responsible for roughly 50% of human cancers due to mutation or inactivation of p53 [38]. By deubiquitinating MDM2 and p53, as well as

other known USP7 substrates, such as tumor suppressors PTEN and FOXO, DNA regulators DNMT1 and UHRF1, and viral E3 ligase ICP0, USP7 plays a role in pathways involved in viral infection, cell signaling, proliferation, and apoptosis [35–37, 39, 59]. Therefore, tight regulation of USP7 activity is essential to maintaining proper cell signaling, and the up–regulation of USP7 appears to be associated with the progression of many cancers [35, 36, 40–43, 56].

Two research groups have identified small molecule compounds that show some selectivity towards the inhibition of USP7 [45,60]. Chauhan and coworkers have identified a compound series that inhibits cell growth and induces apoptosis in multiple myeloma cell lines by inhibiting USP7 [45]. Colland and coworkers have developed a specific inhibitor of USP7 that induces apoptosis, and shows promise as part of a combinatorial therapy [60]. However, both of these compounds are presumed to target the conserved catalytic domain of USP7 and may have off target effects that have yet to be elucidated.

A High–Throughput Screen performed previously within the Mesecar Lab identified a small molecule inhibitor of USP7. This compound was unable to inhibit the catalytic domain alone, but was able to inhibit the full–length enzyme. The work described within this chapter involves the characterization and synthetic optimization of the original hit compound as well as the identification of the required domains within USP7 that are necessary for inhibition.

6.2 Experimental Procedures

The $USP7_{FL}$ enzyme and the truncations utilized in this chapter were expressed and purified as described in Chapter 3.

6.2.1 High–Throughput Screening (HTS) Assay

Disclaimer: This section was performed in collaboration with Dr. Katie Molland, by a previous post–doc in the Mesecar lab. The data included within this section is the extent of what was provided once the project was passed over.

The activity assay described in Section 3.2.8 was optimized for use as a 5 μL reaction with final concentrations of 2.5 nM USP7 and 170 nM Ub–Rho110. HTS was performed in Nunc 1536–well High Base, black polystyrene microplates. USP7 was diluted in HTS Buffer (50 mM Tris pH 7.5, 5 mM DTT, and 0.1 mg/mL BSA) and Ub–Rho110 was diluted in the standard Substrate Buffer (50 mM Tris pH 7.5, 1 mM EDTA, 100 mM NaCl, and 0.05% CHAPS). The assay was performed as an end point assay by quenching the reaction with 0.5 M acetic acid. USP7 was screened against the 25,000 ChemBridge library that was cherry–picked to include compounds composed of drug–like scaffolds, and was also used to screen for USP17 inhibitors as described in Chapter 5. The screen was automated with the use of a ThermoScientific Multidrop Combi reagent dispenser (for addition of the enzyme and substrate) and a BiomekFX liquid handling system (for addition of the compounds). The general scheme of the screen is depicted in Figure 5.1. The screen was performed as follows: 4 μL of 3.13 μM USP7 was dispensed into the plate followed by 20 nL from the 10 mM stocks of the compounds dispensed by a VP Scientific 384–well pin–tool attached to a BiomekFX liquid handling system. The reaction was initiated with 1 μL of 850 nM working concentration of Ub–Rho110, and allowed to incubate for 20 minutes, at which point 1 μL of 0.5 M Acetic Acid was added to quench the reaction. The end–point fluorescence was then read using a Beckman–Coulter DTX880 plate reader at an excitation λ of 485 nm and emission λ of 528 nm. Each plate contained two columns (64 wells) of USP7 without compound present and two columns of substrate alone. The entire screen was performed in singlet at room temperature.

6.2.2 Structure Activity Relationships (SAR) Determination

Preparation of Compounds for SAR Determination

All compounds utilized for the development of the SAR surrounding hit compound **4** (see below) were synthesized at Purdue University by Dr. Antonella Pepe in collaboration with the Computational and Medicinal Chemistry Shared Resource through the Purdue Center for Cancer Research. A total of 23 compounds were synthesized and tested for inhibition of $USP7_{FL}$ (Figure 6.1). The lyophilized compounds were resuspended in 100% DMSO to a stock concentration of 40 mM and stored at -20 °C. The methods used to synthesize these compounds will be published elsewhere.

Inhibition of SAR compounds

The percent inhibition of USP7 at 100 μM of each compound was determined prior to the determination of IC_{50} values to conserve substrate. The final concentration of substrate was held constant at 200 nM and $USP7_{FL}$ was held constant at a final concentration of 1 nM. From the 40 mM stock, each compound was diluted to a working concentration of 3 mM in 100% DMSO. The assay was performed as follows: 1 μL of the working stock of compound was added to a Costar 96 half–volume black plate to which 15 μL of $USP7_{FL}$ was added. Plates were gently mixed and incubated at room temperature for five minutes. To initiate the reaction, 14 μL of Ub–Rho110 was added. Each assay was measured in triplicate. A negative control of Ub–Rho110 alone was measured to evaluate the background rate. Control reactions of $USP7_{FL}$ without compound (DMSO only) were included to measure the rate of the uninhibited USP7 reaction. All reactions contained a final concentration of 3% DMSO. The activity of USP7 was measured as described in Chapter 3 and the % inhibition was calculated using Equation (5.3).

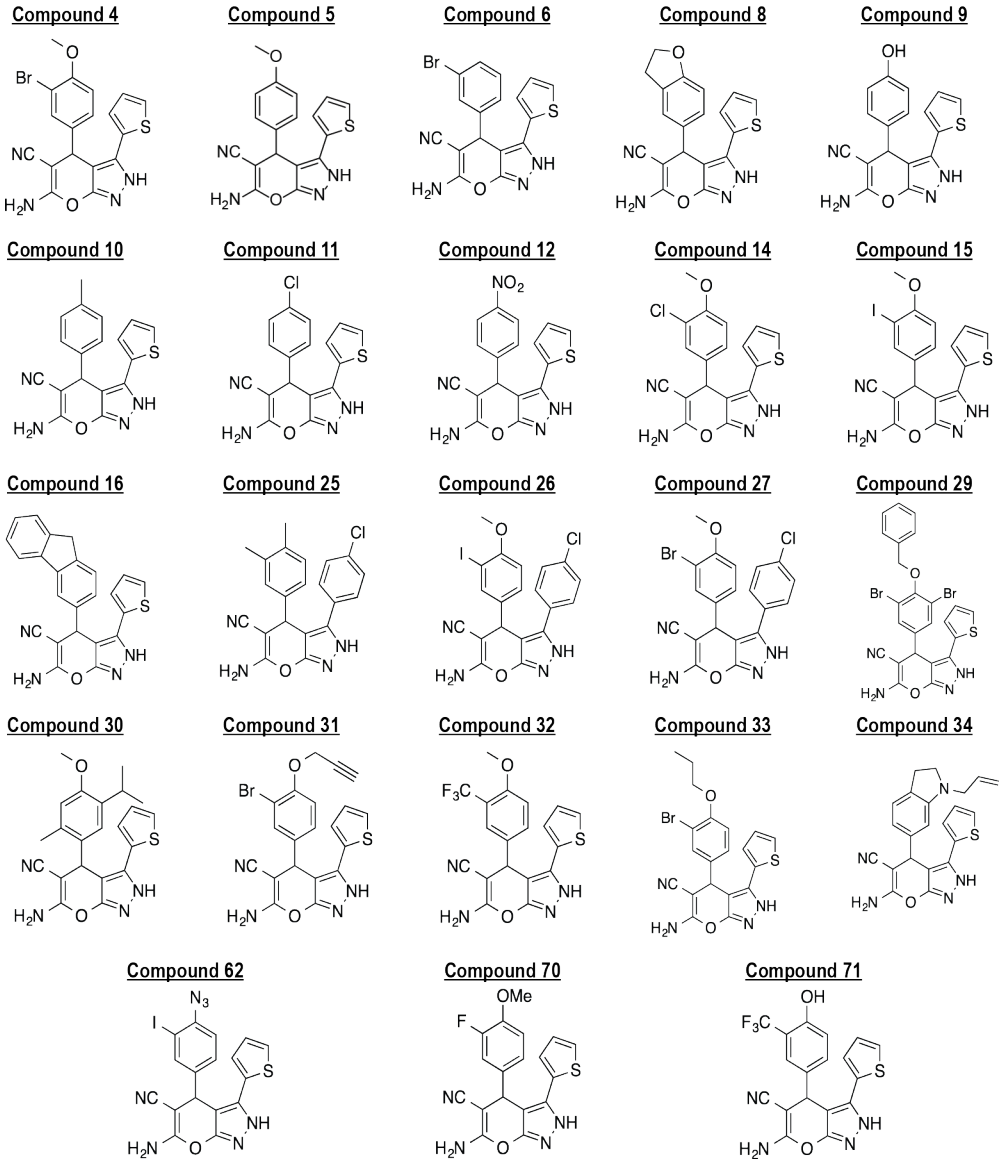


Fig. 6.1. **Compounds Used to Determine SARs with USP7.** Chemical structures of the 23 compounds synthesized to develop the structure activity relationship with USP7 are shown. Each of these compounds was resuspended in 100% DMSO to a stock concentration of 40 mM and then stored at -20 °C.

Determination of IC_{50} values

IC_{50} values were determined for compounds that showed % Inhibition values ≥ 40 %. To determine the IC_{50} values, 11 two–fold serial dilutions of each compound were prepared from the 40 mM stock to working stock concentrations of 6 mM to 0.006 mM in 100% DMSO. The working stocks were added to a Costar 96 half–volume black plate (1 μL) and a final concentration of 1 nM $USP7_{FL}$ in Assay Buffer (15 μL of working stock) was added to each well. The plate was gently mixed and incubated at room temperature for five minutes. To initiate the reaction, a final concentration of 200 nM Ub–Rho110 (14 μL of working stock) was added to each well. The background rate was measured by omitting the enzyme from the reaction. The initial slopes of the progress curves were measured, and the % inhibition was calculated with Equation (5.3). The resulting percentages were plotted in terms of inhibitor concentration. The IC_{50} value was determined from each curve by fitting the data to Equation (5.4) using non–linear regression analysis with SigmaPlot (v13:Systat Software Inc.). For compounds **4**, **15**, and **32**, an IC_{50} value was determined using the same procedure for USP7 truncations $USP7_{CD}$, $USP7_{CD}$–H1–5, $USP7_{CD}$–H4–5 and TRAF–$USP7_{CD}$. The final concentrations of each were 200 nM for $USP7_{CD}$ and TRAF–$USP7_{CD}$, and 1 nM for $USP7_{CD}$–H1–5 and $USP7_{CD}$–H4–5.

6.2.3 Di–Ubiquitin Chain Cleavage Inhibition Assay

Inhibition of $USP7_{FL}$ by compound **4** was evaluated with a more native substrate, Lys^{48} di–ubiquitin (Boston Biochem). This substrate is a more "native" substrate because it contains an iso–peptdie linkage that is normally hydrolyzed by USPs. The reaction was performed with four concentrations of compound **4**: 1.9, 3.9, 39, and 100 μM. Each reaction contained 5 nM $USP7_{FL}$ and 4.2 μM Lys^{48} di–ubiquitin. A negative control was included that lacked $USP7_{FL}$. The positive control contained uninhibited $USP7_{FL}$ and Lys^{48} di–ubiquitin. $USP7_{FL}$ and the compounds were mixed, and incubated at room temperature for 10 minutes prior to initiating the reaction with

the Lys^{48} di–ubiquitin substrate. All reactions were performed at 25 °C in Assay Buffer (Section 2.3.12) for 18 hours. Reactions were quenched with the addition of NuPAGE LDS sample buffer (Invitrogen) supplemented with 5 mM fresh DTT. Samples were analyzed by SDS-PAGE using NuPAGE Novex 4-12% BisTris mini gels (Invitrogen), and quantitatively analyzed using the software AlphaView from ProteinSimple.

6.2.4 Mode of Inhibition Determination

The mode by which compounds **4**, **15**, and **32** inhibit $USP7_{FL}$ was evaluated by varying the compound concentration with fixed–variable concentrations of Ub–Rho110 substrate. Eight two–fold serial dilutions of Ub–Rho110 were made in Substrate Buffer to obtain a working concentration range of 0.08 μM to 10.7 μM. Five two–fold serial dilutions of each compound were made in 100% DMSO to obtain a working concentration range from 46.8 μM to 750 μM. $USP7_{FL}$ was diluted in Assay Buffer to a working concentration of 2 nM. Each assay was performed in 30 μL reaction volumes, in triplicate, in Corning 96 half–well black plates. First, 1 μL of the compound and 15 μL of $USP7_{FL}$ was added to each well, and incubated for five minutes with gentle shaking. To initiate the reaction, 14 μL of Ub–Rho110 was added to the wells, mixed for 10 seconds, and the plate was read for 30 minutes with an excitation λ of 485 nm and emission λ of 528 nm at 25 °C. The uninhibited control contained 1 μL of DMSO, 14 μL of Ub–Rho110, and 15 μL of $USP7_{FL}$ while the negative control lacked enzyme. The final concentrations of each component is as follows: Ub–Rho110 ranged from 0.04 μM to 5 μM, compounds ranged from 1.56 μM to 25 μM and the $USP7_{FL}$ concentration was held constant at 1 nM. The initial rates were determined and converted to v/[E] using the extinction coefficient of the substrate as determined and described in Section 2.3.12. The resulting v/[E] values were plotted as a function of Ub–Rho110 concentration, and the data were fit to four different equations representing Competitive, Non–Competitive, Uncompetitive, and Mixed

inhibition. These equations are standard within the kinetics module of SigmaPlot (v13: Systat Software Inc.). Statistical analysis was performed by comparison of the AIC_C values that were determined from the fits to all four models for each compound tested. The Non–Competitive inhibition equation,

$$\frac{v}{V_{\max}} = \frac{[S]}{K_m \dfrac{\left(1+\dfrac{[I]}{K_I}\right)}{\left(1+\dfrac{\beta[I]}{K_I}\right)} + [S]\dfrac{\left(1+\dfrac{[I]}{K_I}\right)}{\left(1+\dfrac{\beta[I]}{\alpha K_I}\right)}} \tag{6.1}$$

was found to best–fit the data.

6.2.5 Determination of the Steady–State Kinetic Parameters of the TRAF–USP7$_{CD}$ Construct

The steady–state kinetic paramters, K_m and k_{cat} of the TRAF–USP7$_{CD}$ truncation were determined using Ub–Rho110 as a substrate. The enzyme was diluted to a working concentration of 400 nM in Assay Buffer. Twelve, two–fold serial dilutions of Ub–Rho110 were made ranging from 40 μM to 0.04 μM in Substrate Buffer. Assays were performed in 30 μL reaction volumes and in triplicate. The background rates were measured by pipetting 15 μL of Ub–Rho110 at each concentration into a 96–well Corning Costar black half–volume plate, and measuring the rate of change in fluorescence with a BioTek Synergy H1 plate reader. Fluorescence measurements were made every 10 seconds at an excitation λ of 360 nm and emission λ of 460 nm over a time period of two minutes and at 25 °C. Following the measurement of the background rate, 15 μL of the working stock of TRAF–USP7$_{CD}$ was added to the plate, the plate was mixed for 10 seconds, and the assay was measured under the same conditions for 30 minutes total. The final concentration of TRAF–USP7$_{CD}$ was 200 nM and the final concentration range for Ub–Rho110 ranged from 20 μM to 0.01

μM. The initial slope of each reaction in artificial fluorescence units (AFU) per unit time (AFU/sec) was converted to the amount of product produced per unit time (μM/sec) using the extinction coefficient of the substrate. This value of μM/sec was then converted to the rate of reaction in terms of enzyme concentration (V/[E]) by dividing by the enzyme concentration. The data were then analyzed with the kinetics module of SigmaPlot (v13: Systat Software Inc.) by fitting the resulting curves to the Michealis–Menten equation (Equation (2.1)) to determine the k_{cat} and $\mathrm{K_m}$.

6.2.6 Isothermal Calorimetry (ITC) Analysis of Compound Binding

Compound **32** was used for separate ITC analysis with $\mathrm{USP7_{CD}}$ and H1–5 truncations. Measurements were taken using a Malvern MicroCal ITC200 instrument. Initial tests were performed to determine the highest concentration at which compound **32** remained soluble in Analytical Buffer (50 mM HEPES pH 7.5, 3 mM TCEP, and 100 mM NaCl). The 40 mM stock of compound **32** in 100% DMSO was serially diluted in two–fold increments in Analytical Buffer from 200 μM down to 6.25 μM while maintaining 3, 2, 1, 0.5, or 0.25% DMSO. The A_{600} of each sample was measured in comparison to a blank well containing Analytical Buffer and the appropriate percentage of DMSO to determine which well had the least amount of precipitant. The highest concentration of compound **32** at the lowest percentage of DMSO that was found to lack measurable precipitation was determined to be 25 μM and 0.5% DMSO. Due to the lack of solubility of the compound, reverse ITC was performed by diluting the protein into the compound. H1–5 and $\mathrm{USP7_{CD}}$ truncations were dialyzed into Analytical Buffer overnight at 4 °C. The following morning, the samples were concentrated to 515 μM (32 mg/mL) for H1–5 and 736 μM (30 mg/mL) for $\mathrm{USP7_{CD}}$ using spin concentrators with 10,000 MWCO and diluted in Analytical Buffer containing 0.05% DMSO. Both truncations were over 20 times greater in concentration than that of compound **32**. A volume of 200 μL of 25 μM compound **32** was added to the cell and the syringe volume of USP7 truncation was 40 μL. The

initial injection volume was 0.4 μL to remove any bubbles from within the system, and the remaining injections contained 1.8 μL of the USP7 truncation for a total of 22 injections. A three minute delay was programmed between each injection to allow the signal to return to baseline. All reactions were run at 25 °C. The spin speed was set to 1000 rpm with a filter period of five minutes. For each truncation, blank injections were run utilizing the same protocol by injecting the USP7 truncation into Analytical buffer containing 0.5% DMSO. Data was exported into NitPic and the heat from the blank runs were subtracted. The results were analyzed for any measurable heat.

6.3 Results

6.3.1 HTS Hit Identification and Validation

HTS was performed with $USP7_{FL}$ against the cherry-picked Chembridge library containing 25,000 compounds of drug–like scaffolds utilizing Ub–Rho110 as the substrate. The *Z'*-factor's for the screen were calculated as described in Section 5.2.1 and they ranged from 0.75-0.81. A summary of the screen is depicted in Figure 6.2. The percent inhibition threshold was set to 40% to maintain a hit rate of less than 0.1%. Virtual screening to evaluate the chemical properties of each hit, and identify other potential protein targets was performed by Dr. Sergey Savinov. With this inhibition threshold and virtual screening by Dr. Sergey Savinov, 12 compounds were identified as potential hits. Of the 12 compounds, nine were commercially available. From these nine compounds, only four had measurable inhibition towards $USP7_{FL}$. The best of which was compound **4** (Figure 6.3A) which has an IC_{50} value in the single digit micromolar range, 4.1 $\pm$ 0.5 μM and maximum % inhibition of 98 $\pm$ 4 % (Figure 6.3B).

Inhibition of $USP7_{FL}$ by compound **4** was further evaluated with the more physiological substrate Lys^{48} di–ubiquitin. Without compound **4** present in the reaction mixture, roughly 75% of the di–ubiquitin substrate is cleaved (Figure 6.3C lane 2). With increasing concentrations of compound **4**, a steady decrease in the amount of

mono–ubiquitin formed can be observed (Figure 6.3C lanes 3–6). These results indicate that compound **4**, is a bona fide inhibitor of $USP7_{FL}$ and that Ub–Rho110 and Lys^{48} di–ubiquitin can be used as substrates to determine inhibition.

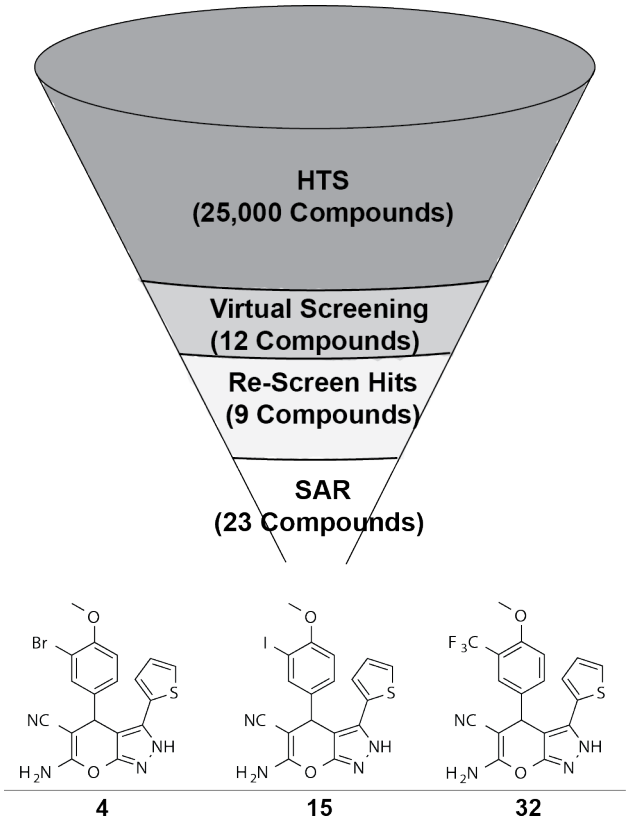


Fig. 6.2. **HTS Hit Validation. A.** A cherry–picked ChemBridge library of 25,000 compounds composed of drug–like scaffolds was screened. Tweleve hits were identified to have 40% inhibition or more. Of the 12 identified compounds, nine were commercially available and further analyzed. Of these nine, four had measurable and reproducible inhibition. SAR was developed using 23 compounds, of which a panel of three was used for the characterization of the inhibition of USP7, compound **4**, **15**, and **32**.

Specificity of compound 4 for $USP7_{FL}$

A small panel of DUB enzymes were tested for inhibition by compound **4** to assess its selectivity for USP7. Both the IC_{50} values and maximum % inhibition were

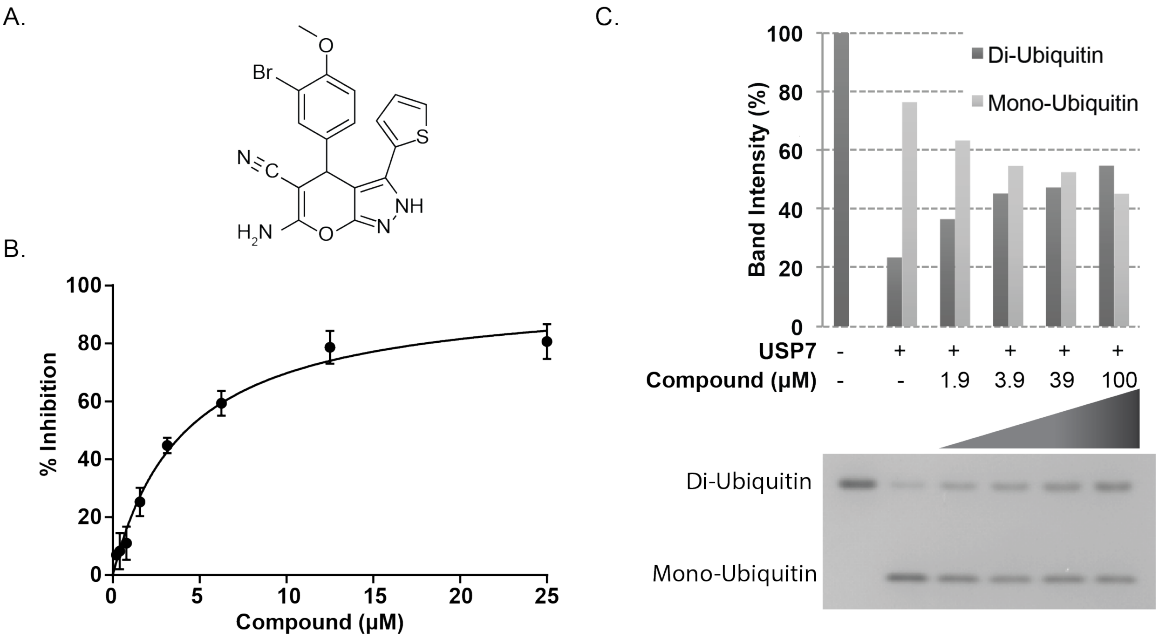


Fig. 6.3. **Kinetic Analysis of HTS Hit Compound 4. A.** Chemical structure of HTS hit, compound **4**. **B.** IC_{50} analysis of compound **4** with 1 nM $USP7_{FL}$ and 200 nM Ub–Rho110. **C.** Cleavage of Lys^{48} di–ubiquitin substrate in the presence of increasing concentrations of compound **4**, as shown. Error bars, SD.

analyzed for compound **4** with two members of the USP family, USP17 and USP28, as well as 2 members of the viral DUB family, SARS–PLpro and MHV–PLP2. The IC_{50} values for compound **4** with these enzymes were all >100 μM with low maximum % inhibition values. The results of this small panel suggest that compound **4** is selective towards USP7, however it would be desirable to test compound **4** against a larger panel of DUBs.

6.3.2 Structure Activity Relationships with the Compound 4 Scaffold

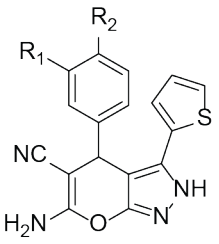
A summary of the compounds analyzed and their corresponding IC_{50} and % inhibition values is shown in Figure 6.4.

Compounds must maintain both R_1 and R_2 substituents on the phenyl ring

The initial round of synthesis focused on determining the importance of the meta–bromine group at R_1 and para–methoxy group at R_2 on the phenyl ring of compound **4**. Replacing either the meta–bromine or the para–methoxy groups with hydrogen (compounds **5** and **6**) abolished inhibitory activity. Furthermore, replacing the meta–bromine group with hydrogen and reducing the size of the para–methoxy group to either a hydroxyl group (compound **9**) or a methyl group (compound **10**) also abolished inhibitory activity.

SAR reveals that inhibition of $\mathrm{USP7_{FL}}$ depends on size and electronegativity of the halogen substituent

The next few rounds of synthesis analyzed the importance in the size and electronegativity of the meta–bromine substituent at R_1. When the bromine substituent was replaced with a fluorine (compound **70**) or chlorine (compound **14**), smaller and more electronegative substituents, the IC_{50} values increased ($IC_{50} = 7 \pm 2$ μM for both). However, if the bromine substituent was replaced with iodine (compound **15**), a larger less electronegative substituent, the IC_{50} value decreased ($IC_{50} = 1.8 \pm 0.1$ μM) indicating tighter binding. Interestingly, when the bromine was replaced with trifluoromethyl (compound **32**), a larger more electronegative substituent, the IC_{50} value decreased further into the nanomolar range ($IC_{50} = 0.75 \pm 0.06$ μM). Figure 6.5 displays the IC_{50} curves for compounds **4**, **14**, **15**, and **32**. Furthermore, when the R_1 bromine was replaced with a trifluoromethyl group and the R_2 group was replaced with a hydroxyl group (compound **71**) an increase in IC_{50} value was again observed, emphasizing the importance of the methoxy group at the R_2 position.



Compound	R_1	R_2	IC_{50} (μM)	Max % Inhibition
4	Br	OCH_3	4.1 ± 0.5	98 ± 4
5	H	OCH_3	> 100	-
6	Br	H	> 100	-
9	H	OH	> 100	-
10	H	CH_3	> 100	-
14	Cl	OCH_3	7 ± 2	91 ± 9
15	I	OCH_3	1.8 ± 0.1	93 ± 2
32	CF_3	OCH_3	0.75 ± 0.06	98 ± 2
30	$CH(CH_3)_2$	OCH_3	1.4 ± 0.2	80 ± 2
70	F	OCH_3	7 ± 2	33 ± 2
71	CF_3	OH	11 ± 0.7	96 ± 2

Fig. 6.4. **SAR Trends Varying in the Phenyl Ring Substituents R_1 and R_2** The effects of varying the subsituents of the phenyl ring of compound **4** on the inhibition of $USP7_{FL}$.

6.3.3 Mode of inhibition of most potent compounds (4, 15, and 32)

The mode of inhibition was determined to be that of Partial Non–Competitive inhibition for this compound series (**4**, **15**, and **32**, Figure 6.6 A). The scheme for Non–Competitive inhibition is shown in Figure 6.7 where the K_I value is the measure of affinity of the compound for both the free enzyme (E) and the enzyme–substrate

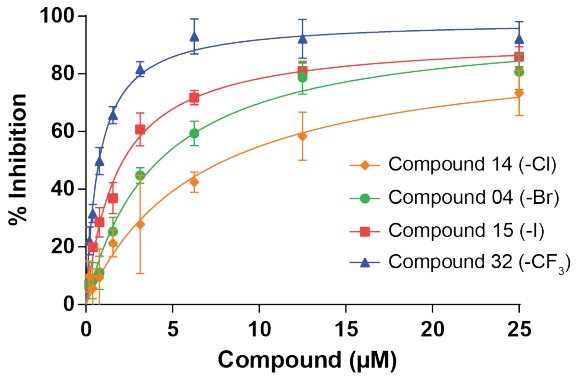


Fig. 6.5. **IC_{50} of Halogen SAR Trend.** IC_{50} curves of $USP7_{FL}$ with compounds **14**, **4**, **15**, and **32**. Assays were composed of 1 nM $USP7_{FL}$, 200 nM Ub–Rho110, and 11 serial dilutions of the compounds ranging from 0.01–25 μM.

complex (ES). As these compounds do not inhibit 100% of the activity of the enzyme, a β value other than one exists which reflects the magnitude by which the k_{cat} value is decreased as compared to the uninhibited enzyme. The K_I and β values for each compound are shown in Table 6.1. The data revealed that the maximum % inhibition increases from compound **4** to **32**, a trend that is reflected in the calculated β values. With Non–Competitive inhibitors the IC_{50} should be equal to the K_I value. While the increase in potency observed with the IC_{50} values are reflected in the K_I values, they are not equal, but the same increase in potency from compound **4** to **32** is observed. The significant decrease in the k_{cat} values observed with increasing concentration of the compounds, and lack of large variation in the K_m values (Figure 6.6 B) suggest all three compounds are V–type allosteric inhibitors (inhibit by decreasing the k_{cat}) and not K–type inhibitors (inhibit by increasing the K_m) of USP7. A typical definition of a Non–Competitive inhibitor is that the inhibitor binds to a separate site on an enzyme that is distinct from the active site such that the inhibitor does not alter the binding affinity of the substrate to the enzyme. Rather the inhibitor decreases the rate of catalysis of the enzyme.

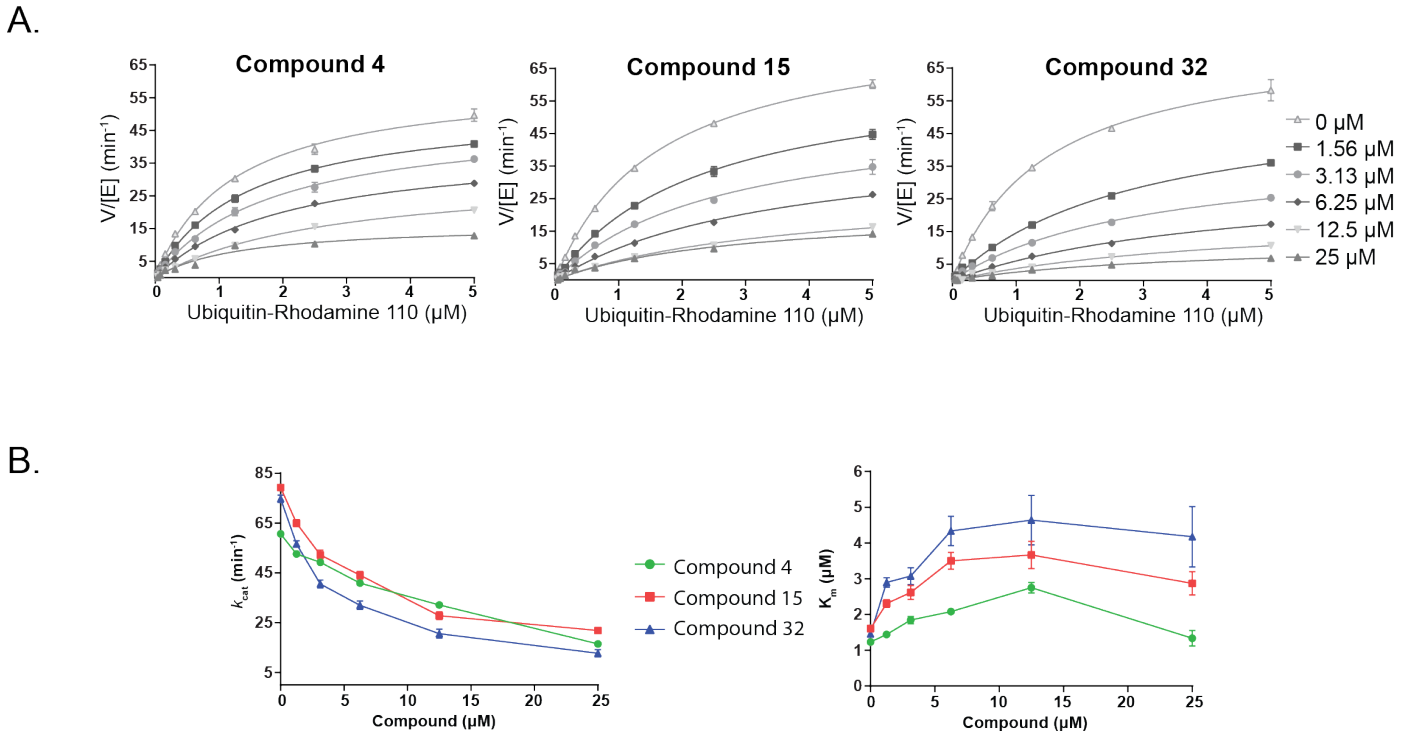


Fig. 6.6. **Mode of Inhibition of compounds 4, 16, and 32. A.** The mode of inhibition was determined for compounds **4**, **15** and **32** with $USP7_{FL}$. Assays were performed by varying the concentration of Ub–Rho110 from 0.04–5 μM with fixed–variable concentrations of the compounds from 1.56–25 μM. An uninhibited $USP7_{FL}$ control was also measured. Initial RFU/min were measured for each curve and were converted using the determined extinction coefficient of Ub–Rho110 to V/[E]. Resulting values were plotted against the concentration of Ub–Rho110. **B.** Each curve in A. was fit to the Michealis–Menten equation to determine the k_{cat} and K_m values. The resulting values were plotted in terms of the compound concentration. Error bars, SD.

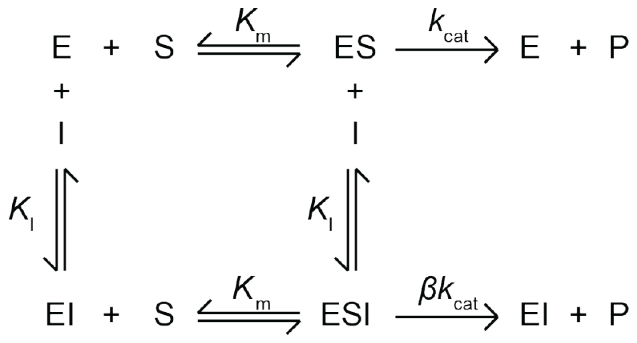


Fig. 6.7. **Scheme for Partial Non–Competitive Inhibition.** Expansion of the typical Non–Competitive enzyme kinetic scheme to include partial inhibition. The Partial Non–Competitive inhibition model includes the kinetic parameter, K_I, and the kinetic modulator parameter β, which modulates k_{cat}.

Table 6.1
Partial Non–Competitive inhibition parameters for compounds **4**, **15**, and **32**

Compound	β	K_I
4	0.209 ± 0.007	4.3 ± 0.2
15	0.150 ± 0.006	2.48 ± 0.09
32	0.072 ± 0.005	1.60 ± 0.06

6.3.4 IC_{50} panel of compound 4, 15, and 32 with USP7 truncations

In an effort to determine where the alternative binding site on the enzyme for the Partial Non–Competitive inhibitors is, IC_{50} values were determined for all three compounds against four different truncations of USP7: $USP7_{CD}$, $USP7_{CD}$–H1–5, $USP7_{CD}$–H4–5, and TRAF–$USP7_{CD}$ (Figure 6.8). In conjunction with the mode of inhibition analysis, the lack of inhibition of the $USP7_{CD}$ (Figure 6.8) suggests that these compounds do not bind to the conserved catalytic domain. Rather, these compounds require the other domains of USP7 to produce inhibition. Interestingly, the IC_{50} values could only be measured for $USP7_{CD}$–H1–5(Figure 6.8: compound **4** = 11

$\pm$ 5 μM, compound **15** = 6 $\pm$ 2 μM and compound **32** = 3 $\pm$ 2 μM). These IC_{50} values are significantly larger than those measured for $USP7_{FL}$ suggesting that the TRAF domain is also required (Figure 6.4). Taken together, these results suggest that the HUBL domains are essential for inhibition, specifically HUBL domains 1, 2, and 3 as no inhibition could be measured with $USP7_{CD}$–H4–5. Furthermore, the TRAF domain appears to be important but not absolutely essential for inhibition.

6.3.5 Role of the TRAF domain in inhibition of $USP7_{FL}$ by compounds 4, 15, and 32

To further investigate the role of the TRAF domain in the inhibition of USP7, the steady–state kinetic parameters, K_m and k_{cat}, values were determined for TRAF–$USP7_{CD}$ using Ub–Rho110. THe values were found to be 25 $\pm$ 1 μM and 0.174 $\pm$ 0.005 s^{-1} respectively. The k_{cat} value of TRAF–$USP7_{CD}$ is roughly 4 fold higher than the k_{cat} value of $USP7_{CD}$ and is 7 fold lower than the k_{cat} value of $USP7_{FL}$. Interestingly, a large difference is observed between the K_m values, where the TRAF–$USP7_{CD}$ K_m value is $\sim$ 16 fold higher than that of the K_m value for both $USP7_{FL}$ and $USP7_{CD}$.

6.3.6 Isothermal Calorimetry Analysis of Compound 32

Isothermal calorimetry (ITC) was utilized to further evaluate the required domains of USP7 for inhibition by these compounds. The most potent inhibitor, compound **32**, was utilized. However, due to solubility issues, the ITC experiments were performed in reverse by diluting a high concentration of USP7 truncation into compound **32**. H1–5 and $USP7_{CD}$ truncations were injected into 25 μM of compound **32** at 515 μM and 736 μM, respectively. The background heat was measured by injecting the USP7 truncation into Analytic Buffer containing 0.5% DMSO. Neither H1–5 or $USP7_{CD}$ injections generated heat greater than the background levels. These results may suggest one of two results. First, the interaction between compound **32** and USP7 is not

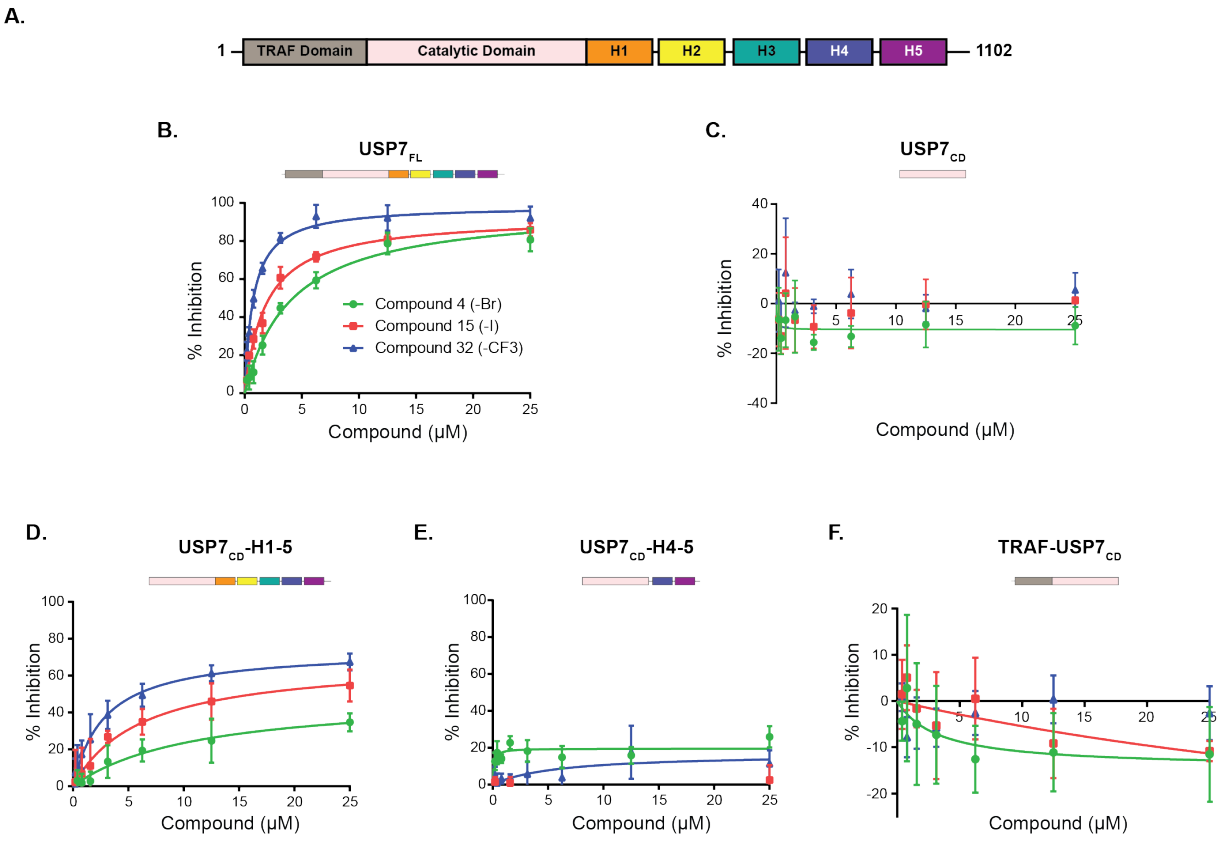


Fig. 6.8. **Inhibition of Different USP7 Truncations by Designed Compounds. A.** Schematic of the 7 different domains of USP7. N–terminal TRAF domain (grey), catalytic domain (pink), five ubiquitin–like domains (rainbow). IC_{50} curves for compound **4** (green circles), **15** (red squares), and **32** (blue triangles) were determined with 200 nM Ub–Rho110. Compound concentration ranged from 0.01 to 25 μM. Enzyme concentrations were held constant at 1 nM for $USP7_{FL}$, $USP7_{CD}$–H1–5, and $USP7_{CD}$–H4–5 and 200 nM for $USP7_{CD}$ and TRAF–$USP7_{CD}$. IC_{50} curves of **B.** $USP7_{FL}$, **C.** $USP7_{CD}$, **D.** $USP7_{CD}$–H1–5, **E.** $USP7_{CD}$–H4–5, **F.** TRAF–$USP7_{CD}$. Data in panels B and D were fit to Equation (5.4) to determine IC_{50} values. Error bars, SD.

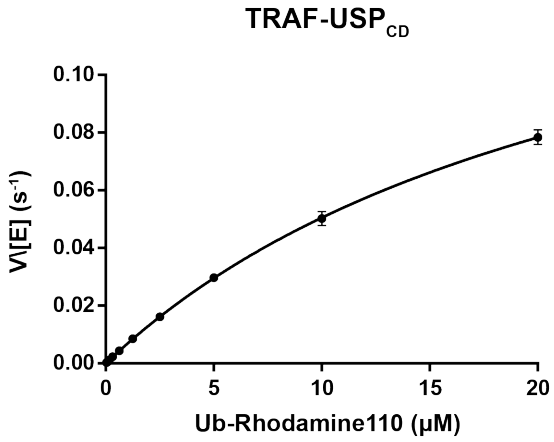


Fig. 6.9. **Interaction of Ub–Rho110 with TRAF–USP7$_{CD}$.** The rates of hydrolysis of Ub–Rho110 by the TRAF–USP7$_{CD}$ are shown and are normalized to the total TRAF–USP7$_{CD}$ concentration. TRAF–USP7$_{CD}$ concentration is 200 nM. Ub–Rho110 concentrations varied from 0.02 μM to 20 μM. Data were performed in triplicate.

enthalpically driven. Second, compound **32** cannot bind the individual truncations, and requires an interface to be formed between the domains for binding.

6.4 Discussion

The goal of developing inhibitors towards specific members of the USP family mirrors that of the protein kinase family of enzymes. It was long assumed that targeting a specific protein kinase within the family that shares high similarity in both structure and sequence was an impossible task. However, in 2001 the first specific kinase inhibitor, imatinib, was FDA approved for targeting BCR–ABL cancers [67]. The impact of this accomplishment has been widespread since 2001, 27 more drugs that target specific protein kinases (i.e. EGRF, Src, B–raf, etc) have entered the market [67]. Furthermore, this success broadened our horizons with regards to targeting specific enzymes within a large family, including the USPs which share a conserved catalytic domain.

Two groups have successfully identified small molecule inhibitors that are presumably specific against USP7 and these compounds have shown efficacy in cell culture [45, 60]. However, both of these molecules target the conserved catalytic domain and while both molecules have been shown to be specific towards USP7 *in vitro*, off target effects may yet arise when further tested *in vivo*.

Several groups have reported significantly lower activity of $USP7_{CD}$ than $USP7_{FL}$ [20, 34]. In our hands, the difference in activity between $USP7_{CD}$ and $USP7_{FL}$ was 30 fold. Faesen and coworkers elucidated the requirement of HUBL domains 4 and 5 in the intramolecular activation of USP7 by a proposed interaction between the C–terminal loop of HUBL 5 and the switching loop within $USP7_{CD}$. However, the involvement of HUBL domains 1, 2, or 3 in the intramolecular activation was not defined. The work described in Chapter 4 elucidates the role of HUBL domains 1, 2, and 3 in the intramolecular activation of USP7 as required for the regulation of HUBL domains 4 and 5. Furthermore, it has been observed that the TRAF domain does not have a role in the intramolecular activation of USP7 and is dispensable for activity [20]. Understanding how USP7 activity is internally regulated by the HUBL domains has aided in the development of USP7 inhibitors discussed within this chapter.

Compound **4** (Figure 6.3 A) was identified from the HTS screen and it has an IC_{50} value of 4.1 μM. Compound **4** was found to be specific towards USP7 using a small panel of DUB enzymes. Modifications to the meta–bromine and para–methoxy substituents of the phenyl ring elucidated the requirement for both substituents to be present for inhibition. When the para–methoxy group was reduced in size or electronegativity, inhibition was lost. Potency was increased with an increase in size and electronegativity at the meta position. While these results strongly suggest that the pocket the phenyl group occupies is large and prefers a highly electronegative group, without a structure of USP7 bound to these compounds, we cannot further discuss the importance of these elucidated features from a structural standpoint.

The compounds developed in this work were determined to inhibit USP7 by a Partial Non–Competitive mode of inhibition (Figure 6.6) which suggests that they

do not bind to the conserved catalytic active site. Upon further analysis, it was discovered that the inhibitors could not inhibit $USP7_{CD}$ alone. Strikingly, when the compounds were tested against $USP7_{CD}$–H1–5 and $USP7_{CD}$–H4–5, it was found that HUBL domains 1–3 were essential for inhibition. These observations suggest that these compounds may bind to an interface on HUBL domains 1–3 and thereby disrupts the stabilizing influence these domains have on the intramolecular activation of USP7 by HUBL domains 4 and 5 (Chapter 4).

Interestingly, the IC_{50} values determined with $USP7_{CD}$–H1–5 were significantly weaker than the values determined with $USP7_{FL}$, however no inhibition could be measured with the TRAF–$USP7_{CD}$ construct. These results suggest that the TRAF domain may have a role in the activity of USP7 after all, and that the TRAF domain acts in concert with HUBL domains 1–5 to produce maximum inhibition. We hypothesize that the significantly higher K_m value associated with TRAF–$USP7_{CD}$ indicates that the TRAF domain may bind Ub–Rho110, and may require the HUBL domains for proper positioning with the catalytic domain. Without the HUBL domains, the binding of Ub–Rho110 increases the K_m value by reducing the amount of free substrate that can enter the active site. An interaction between the TRAF domain and the HUBL domains has yet to be described, however the results presented here suggest further investigation is necessary. The increase in the k_{cat} value from $USP7_{CD}$ alone suggests that the TRAF domain, when aligned correctly aids in stabilizing Ub–Rho110 with the catalytic domain.

USP7 is an important drug target for many types of cancers due to its up-regulation and the signaling pathways its involved in. The compounds described here appear to require the HUBL domains and the TRAF domain for full inhibition. Furthermore, these results suggest that the compounds disrupt the intramolecular activation of USP7, imposed by the HUBL and TRAF domains. By targeting the interactions of all the domains within USP7 and their function in the intramolecular activation of USP7 these compounds show great promise for further development as USP7 specific inhibitors. This unique approach would allow for reduced activity of

USP7 but not complete inhibition, which could lessen the effects of full inhibition of USP7 on other important pathways in the cell. Furthermore, drugs developed to target the HUBL domains would be more specific to USP7 due to the unique architecture of the USP7 domains, as opposed to targeting the highly-conserved structure of the catalytic domain.

CHAPTER 7. SUMMARY

Many members of the Ubiquitin Specific Protease (USP) family play a role in the regulation of cell–cycle progression. As a result, their observed upregulation in many disease states, predominantly cancer, suggests that individual members of this family of enzymes would make ideal drug targets for anti–cancer therapeutic development. The USP family members are composed of a conserved catalytic domain which is flanked by a variety of domains important for substrate recognition, regulation, and catalysis. It is the composition of these flanking domains that make each USP unique. While the conserved catalytic domain of the USPs share high structural similarities, two regions within the catalytic domain share high sequence similarity as well, the Cys and His boxes. The catalytic triad is composed of a cysteine, histidine, and aspartate residues and are contained with these sequence conserved boxes. These similarities make the task of inhibiting specific members of the USP family substantially more difficult. Therefore, thorough enzymatic and biochemical characterization of the USP of interest is required, specifically the characterization of the role of the flanking domains that make each USP unique. Along these lines, we report here the enzymatic and biochemical characterization of two USP family members, USP17 and USP7 are reported in this dissertation.

USP17 is an important anti–cancer target for its role in perturbing both the Ras and CDC25A signaling pathways in breast and prostate cancers. As such, the expression, refolding, and purification of USP17 was pursued and achieved. Due to its insolubility when expressed in *E. coli* expression systems, the recombinant expression and purification of USP17 had not previously been reported. Furthermore, while numerous USPs have successfully been expressed and purified using the baculovirus expression system, it was found that expression of USP17 within the baculovirus system results in its significant degradation and low yields. A methodology for the

refolding of USP17 from *E. coli* inclusion body was developed and results in ∼4 mg/L of pure USP17.

Having the ability to purify milligram quantities of USP17 allowed for the enzymatic and biochemical characterization of USP17. The enzymatic characterization of USP17 included defining the steady–state kinetic parameters, k_{cat} and K_m, with fluorogenic substrates Ub–Rho110 and Ub–AMC and determination of USP17's ubiquitin–chain linkage specificity. Comparison of the determined steady–state kinetic values with the values of other characterized USPs, it was discovered that USP17 has significantly higher enzymatic activity than many members of the USP family. Furthermore, as is common with USPs, USP17 is very promiscuous with respect to ubiquitin–chain linkage recognition, and could cleave all seven lysine linkages to a certain extent but could not cleave linear ubiquitin chains. Biochemical characterization of USP17 included determination of the oligomeric state, stability, and preferred pH range for activity. The oligomeric state of USP17 was determined using 3 different techniques: Analytical Size Exclusion Chromatography, Analytical Ultracentrifugation, and Dynamic Light Scattering. Combining the results of these techniques suggests that USP17 is a monomer in solution. The preferred pH range for optimal USP17 activity was between pH 7 and 9. As this is the first report of the enzymatic and biochemical characterization of USP17, this work focused on the conserved catalytic domain of USP17, however to develop specific inhibitors towards USP17, a deeper understanding of the C–terminus hyaluronan binding motifs is required.

The USP17 kinetic assay utilizing Ub–Rho110 was optimized for high throughput–screening (HTS) conditions, and a cherry–picked ChemBridge library composed of 25,000 drug–like scaffolds was screened. From this screen and the secondary screening of the DiverSet ChemBridge library, four promising scaffolds were identified. A structure activity relationship (SAR) was developed from the analysis of the compounds within the screens containing one of the four scaffolds. Thirteen compounds containing the identified scaffolds from the ChemBridge Hit–2–Lead library were or-

dered to further develop the SAR, and evaluated for % inhibition and IC_{50} values. Only two of the 13 compounds had measurable IC_{50} values, one of which is ideal to pursue further. The low hit rate of this screen suggests further optimization of the HTS assay, and future screening of more diverse libraries.

USP7/HAUSP is a highly characterized USP that was originally identified for its role in herpesvirus infection, and has since been defined to regulate many tumor suppressor related pathways, including the p53/MDM2 pathway. The upregulation of USP7 is described in a variety of cancers, including prostate, breast, glioma, lung, multiple myeloma, colon, and neuroblastoma [35,36,40–43,56]. The catalytic domain of USP7 has significantly lower activity than the full–length enzyme, and requires intramolecular activation by the C–terminal ubiquitin–like domains (HUBL) 4 and 5. The role of HUBL domains 1–3 in this intramolecular activation, however, was previously ill defined. Through kinetic and biochemical characterization of USP7, a model displaying the significance of HUBL domains 1–3 in the intramolecular activation of USP7 was developed and the role of HUBL domains 1–3 in inhibition by identified compounds was established.

To evaluate the role of the HUBL domains in the intramolecular activation of USP7, nine different USP7 constructs were expressed and purified: $USP7_{FL}$, $USP7_{CD}$, $USP7_{CD}$–H1–5, $USP7_{CD}$–H4–5, TRAF–$USP7_{CD}$, H1–5, H4–5, H1–2, and H1–3. With the exception of $USP7_{FL}$ which was prepared from the baculovirus expression system, each truncation was expressed from *E. coli* and purified using three sequential chromotographic steps. All truncations were purified in milligram quantities of $\geq$ 98% purity.

Using the fluorgenic substrate, Ub–Rho110, the steady–state kinetic parameters for the active constructs ($USP7_{FL}$, $USP7_{CD}$, $USP7_{CD}$–H4–5, and TRAF–$USP7_{CD}$) were determined. The results validated the requirement of HUBL domains 4 and 5 to achieve full activity of USP7. Interestingly, the parameters associated with the TRAF–$USP7_{CD}$ truncation suggests that the TRAF domain interacts with the Ub–Rho110 substrate, a result previously not observed. Further kinetic approaches, as

well as other biochemical and biophysical experiments allowed the assignments of the role of then HUBL domains 1–3 in the intramolecular activation of USP7. Based on our results, we propose a novel rheostat model for the regulation of USP7 by HUBL domains 1–3. HUBL domains 1–3 serve to adjust the level of activity USP7 experiences by stabilizing and regulating the interaction of HUBL domain's 4 and 5 with the catalytic domain.

A promising small molecule inhibitor of USP7 was identified using $USP7_{FL}$ by HTS. In collaboration with the Computational and Medicinal Chemistry core of the Purdue Center for Cancer Research, several rounds of synthesis allowed for the develop a strong SAR surrounding the halogen substituent of the scaffold. The different active truncations of USP7 were tested for inhibition by these compounds, and it was found that no inhibition could be measured without HUBL domains 1–3 present. Fitting with this finding, the mode of inhibition of these compounds was determined to be Non–Competitive, suggesting these inhibitors act via an allosteric mechanism. The allosteric inhibitors identified by this dissertation work are the first of their kind, and will provide a novel chemical tool to the scientific community. Many attempts were made to determine where these compounds bind to USP7 using both kinetic and biophysical (ITC) approaches, however no conclusive results were obtained. It is hypothesized that these compounds bind to an interface formed between HUBL domains 1–3 and the catalytic domain, and, without both components, binding does not occur. Future work on this project will be both structural and biophysical in nature to determine where these compounds bind to USP7 and how these compounds inhibit USP7 activity.

LIST OF REFERENCES

LIST OF REFERENCES

[1] J. M. Fraile, V. Quesada, D. Rodriguez, J. M. P. Freije, and C. Lopez-Otin, "Deubiquitinases in cancer: new functions and therapeutic options," *Oncogene*, vol. 31, no. 19, pp. 2373–2388, 2012.

[2] S. M. B. Nijman, M. P. A. Luna-Vargas, A. Velds, T. R. Brummelkamp, A. M. G. Dirac, T. K. Sixma, and R. Bernards, "A Genomic and Functional Inventory of Deubiquitinating Enzymes," *Cell*, vol. 123, no. 5, pp. 773–786, 2005.

[3] C. M. Pickart and D. Fushman, "Polyubiquitin chains: polymeric protein signals," *Current Opinion in Chemical Biology*, vol. 8, no. 6, pp. 610–616, 2004.

[4] A. Y. Amerik and M. Hochstrasser, "Mechanism and function of deubiquitinating enzymes," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1695, no. 1, pp. 189–207, 2004.

[5] C. M. Pickart and M. J. Eddins, "Ubiquitin: structures, functions, mechanisms," *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, vol. 1695, no. 13, pp. 55–72, 2004.

[6] A. Sarikas, T. Hartmann, and Z.-Q. Pan, "The cullin protein family," *Genome Biology*, vol. 12, no. 4, p. 220, 2011.

[7] A. Williamson, K. E. Wickliffe, B. G. Mellone, L. Song, G. H. Karpen, and M. Rape, "Identification of a physiological E2 module for the human anaphase-promoting complex," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 43, pp. 18213–18218, 2009.

[8] A. K. Al-Hakim, A. Zagorska, L. Chapman, M. Deak, M. Peggie, and D. R. Alessi, "Control of AMPK-related kinases by USP9X and atypical Lys29/Lys33-linked polyubiquitin chains," *Biochemical Journal*, vol. 411, no. 2, pp. 249–260, 2008.

[9] S. M. B. Nijman, T. T. Huang, A. M. G. Dirac, T. R. Brummelkamp, R. M. Kerkhoven, A. D. D'Andrea, and R. Bernards, "The Deubiquitinating Enzyme USP1 Regulates the Fanconi Anemia Pathway," *Molecular Cell*, vol. 17, no. 3, pp. 331–339, 2005.

[10] D. Komander, M. J. Clague, and S. Urbé, "Breaking the chains: structure and function of the deubiquitinases.," *Nature reviews. Molecular cell biology*, vol. 10, pp. 550–63, aug 2009.

[11] A. C. Faesen, M. P. A. Luna-Vargas, P. P. Geurink, M. Clerici, R. Merkx, W. J. van Dijk, D. S. Hameed, F. El Oualid, H. Ovaa, and T. K. Sixma, "The Differential Modulation of USP Activity by Internal Regulatory Domains, Interactors and Eight Ubiquitin Chain Types," *Chemistry & Biology*, vol. 18, no. 12, pp. 1550–1561, 2011.

[12] D. Fushman and K. D. Wilkinson, "Structure and recognition of polyubiquitin chains of different lengths and linkage," *F1000 biology reports*, vol. 3, p. 26, 2011.

[13] J. J. Sacco, J. M. Coulson, M. J. Clague, and S. Urbé, "Emerging roles of deubiquitinases in cancer-associated pathways.," *IUBMB life*, vol. 62, pp. 140–57, feb 2010.

[14] Y. Ye, H. Scheel, K. Hofmann, and D. Komander, "Dissection of USP catalytic domains reveals five common insertion points.," *Molecular bioSystems*, vol. 5, pp. 1797–808, dec 2009.

[15] K. Molland, Q. Zhou, and A. D. Mesecar, "A 2.2 A resolution structure of the USP7 catalytic domain in a new space group elaborates upon structural rearrangements resulting from ubiquitin binding," *Acta Crystallographica Section F*, vol. 70, no. 3, pp. 283–287, 2014.

[16] M. Hu, P. Li, M. Li, W. Li, T. Yao, J.-W. Wu, W. Gu, R. E. Cohen, and Y. Shi, "Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde," *Cell*, vol. 111, no. 7, pp. 1041–1054, 2002.

[17] M. A. Villamil, J. Chen, Q. Liang, and Z. Zhuang, "A Noncanonical Cysteine Protease USP1 Is Activated through Active Site Modulation by USP1-Associated Factor 1," *Biochemistry*, vol. 51, no. 13, pp. 2829–2839, 2012.

[18] G. V. Avvakumov, J. R. Walker, S. Xue, P. J. Finerty Jr., F. Mackenzie, E. M. Newman, and S. Dhe-Paganon, "Amino-terminal dimerization, NRDP1-rhodanese interaction, and inhibited catalytic domain conformation of the ubiquitin-specific protease 8 (USP8)," *Journal of Biological Chemistry*, vol. 281, no. 49, pp. 38061–38070, 2006.

[19] M. Hu, P. Li, L. Song, P. D. Jeffrey, T. A. Chenova, K. D. Wilkinson, R. E. Cohen, and Y. Shi, "Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14," *The EMBO Journal*, vol. 24, no. 21, pp. 3747–3756, 2005.

[20] A. Faesen, A. G. Dirac, A. Shanmugham, H. Ovaa, A. Perrakis, and T. Sixma, "Mechanism of USP7/HAUSP Activation by Its C-Terminal Ubiquitin-like Domain and Allosteric Regulation by GMP-Synthetase," *Molecular Cell*, vol. 44, no. 1, pp. 147–159, 2011.

[21] M. Clerici, M. P. A. Luna-Vargas, A. C. Faesen, and T. K. Sixma, "The DUSP-Ubl domain of USP4 enhances its catalytic efficiency by promoting ubiquitin exchange.," *Nature communications*, vol. 5, p. 5399, 2014.

[22] A. Faesen, M. Luna-Vargas, and T. Sixma, "The role of UBL domains in ubiquitin-specific proteases," *Biochemical Society Transactions*, vol. 40, no. 3, pp. 539–545, 2012.

[23] X. Zhu, R. Ménard, and T. Sulea, "High incidence of ubiquitin-like domains in human ubiquitin-specific proteases," *Proteins: Structure, Function, and Bioinformatics*, vol. 69, no. 1, pp. 1–7, 2007.

[24] F. Colland, "The therapeutic potential of deubiquitinating enzyme inhibitors.," *Biochemical Society transactions*, vol. 38, pp. 137–43, feb 2010.

[25] J. F. Burrows, M. J. McGrattan, A. Rascle, M. Humbert, K. H. Baek, and J. A. Johnston, "DUB-3, a cytokine-inducible deubiquitinating enzyme that blocks proliferation," *Journal of Biological Chemistry*, vol. 279, no. 14, pp. 13993–14000, 2004.

[26] J.-M. Shin, K.-J. Yoo, M.-S. Kim, D. Kim, and K.-H. Baek, "Hyaluronan- and RNA-binding deubiquitinating enzymes of USP17 family members associated with cell viability," *Bmc Genomics*, vol. 7, 2006.

[27] C. McFarlane, A. A. Kelvin, M. de la Vega, U. Govender, C. J. Scott, J. F. Burrows, and J. A. Johnston, "The Deubiquitinating Enzyme USP17 Is Highly Expressed in Tumor Biopsies, Is Cell Cycle Regulated, and Is Required for G(1)-S Progression," *Cancer Research*, vol. 70, no. 8, pp. 3329–3339, 2010.

[28] J. F. Burrows, A. A. Kelvin, C. McFarlane, R. E. Burden, M. J. McGrattan, M. De la Vega, U. Govender, D. J. Quinn, K. Dib, M. Gadina, C. J. Scott, and J. A. Johnston, "USP17 Regulates Ras Activation and Cell Proliferation by Blocking RCE1 Activity," *Journal of Biological Chemistry*, vol. 284, no. 14, pp. 9587–9595, 2009.

[29] Y. Pereg, B. Y. Liu, K. M. O'Rourke, M. Sagolla, A. Dey, L. Komuves, D. M. French, and V. M. Dixit, "Ubiquitin hydrolase Dub3 promotes oncogenic transformation by stabilizing Cdc25A," *Nature Cell Biology*, vol. 12, no. 4, pp. 400–U226, 2010.

[30] J. Jaworski, U. Govender, C. McFarlane, M. de la Vega, M. K. Greene, N. D. Rawlings, J. A. Johnston, C. J. Scott, and J. E. Burrows, "A novel RCE1 isoform is required for H-Ras plasma membrane localization and is regulated by USP17," *Biochemical Journal*, vol. 457, pp. 289–300, 2014.

[31] S. D. Hayes and J. W. Harper, "Cdc25A and Dub3 in a high-stakes balancing act," *Nature Cell Biology*, vol. 12, no. 4, pp. 311–313, 2010.

[32] Y.-T. Chiu, H.-Y. Han, S. C.-L. Leung, H.-F. Yuen, C.-W. Chau, Z. Guo, Y. Qiu, K.-W. Chan, X. Wang, Y.-C. Wong, and M.-T. Ling, "CDC25A Functions as a Novel Ar Corepressor in Prostate Cancer Cells," *Journal of Molecular Biology*, vol. 385, no. 2, pp. 446–456, 2009.

[33] M. Hu, L. Gu, M. Li, P. D. Jeffrey, W. Gu, and Y. Shi, "Structural Basis of Competitive Recognition of p53 and MDM2 by HAUSP/USP7: Implications for the Regulation of the p53MDM2 Pathway," *PLoS Biol*, vol. 4, no. 2, p. e27, 2006.

[34] A. Fernández-Montalván, T. Bouwmeester, G. Joberty, R. Mader, M. Mahnke, B. Pierrat, J.-M. Schlaeppi, S. Worpenberg, and B. Gerhartz, "Biochemical characterization of USP7 reveals post-translational modification sites and structural requirements for substrate processing and subcellular localization," *FEBS Journal*, vol. 274, no. 16, pp. 4256–4270, 2007.

[35] M. S. Song, L. Salmena, A. Carracedo, A. Egia, F. Lo-Coco, J. Teruya-Feldstein, and P. P. Pandolfi, "The deubiquitinylation and localization of PTEN are regulated by a HAUSP-PML network," *Nature*, vol. 455, no. 7214, pp. 813–817, 2008.

[36] A. van der Horst, A. M. M. de Vries-Smits, A. B. Brenkman, M. H. van Triest, N. van den Broek, F. Colland, M. M. Maurice, and B. M. T. Burgering, "FOXO4 transcriptional activity is regulated by monoubiquitination and USP7/HAUSP," *Nat Cell Biol*, vol. 8, no. 10, pp. 1064–1073, 2006.

[37] J. Cheng, H. Yang, J. Fang, L. Ma, R. Gong, P. Wang, Z. Li, and Y. Xu, "Molecular mechanism for USP7-mediated DNMT1 stabilization by acetylation.," *Nature communications*, vol. 6, p. 7023, 2015.

[38] B. Vogelstein, D. Lane, and A. J. Levine, "Surfing the p53 network," *Nature*, vol. 408, no. 6810, pp. 307–310, 2000.

[39] R. Pfoh, I. K. Lacdao, A. A. Georges, A. Capar, H. Zheng, L. Frappier, and V. Saridakis, "Crystal Structure of USP7 Ubiquitin-like Domains with an ICP0 Peptide Reveals a Novel Mechanism Used by Viral and Cellular Proteins to Target USP7," *PLoS Pathog*, vol. 11, no. 6, pp. 1–23, 2015.

[40] J. M. Cummins and B. Vogelstein, "HAUSP is required for p53 destabilization," *Cell Cycle*, vol. 3, no. 6, pp. 689–692, 2004.

[41] N. Kon, Y. Kobayashi, M. Li, C. L. Brooks, T. Ludwig, and W. Gu, "Inactivation of HAUSP in vivo modulates p53 function," *Oncogene*, vol. 29, no. 9, pp. 1270–1279, 2010.

[42] M. Li, C. L. Brooks, N. Kon, and W. Gu, "A Dynamic Role of HAUSP in the p53-Mdm2 Pathway," *Molecular Cell*, vol. 13, no. 6, pp. 879–886, 2004.

[43] E. Meulmeester, M. M. Maurice, C. Boutell, A. F. A. S. Teunisse, H. Ovaa, T. E. Abraham, R. W. Dirks, and A. G. Jochemsen, "Loss of HAUSP-Mediated Deubiquitination Contributes to DNA Damage-Induced Destabilization of Hdmx and Hdm2," *Molecular Cell*, vol. 18, no. 5, pp. 565–576, 2005.

[44] T. Stühmer, M. Chatterjee, M. Hildebrandt, P. Herrmann, H. Gollasch, C. Gerecke, S. Theurich, L. Cigliano, R. A. Manz, P. T. Daniel, K. Bommert, L. T. Vassilev, and R. C. Bargou, "Nongenotoxic activation of the p53 pathway as a therapeutic strategy for multiple myeloma.," *Blood*, vol. 106, pp. 3609–17, nov 2005.

[45] D. Chauhan, Z. Tian, B. Nicholson, K. G. S. Kumar, B. Zhou, R. Carrasco, J. McDermott, C. Leach, M. Fulcinniti, M. Kodrasov, J. Weinstock, W. Kingsbury, T. Hideshima, P. Shah, S. Minvielle, M. Altun, B. Kessler, R. Orlowski, P. Richardson, N. Munshi, and K. Anderson, "A Small Molecule Inhibitor of Ubiquitin-Specific Protease-7 Induces Apoptosis in Multiple Myeloma Cells and Overcomes Bortezomib Resistance," *Cancer Cell*, vol. 22, no. 3, pp. 345–358, 2012.

[46] C. McFarlane, S. McFarlane, I. Paul, K. Arthur, M. Scheaff, K. Kerr, M. Stevenson, D. A. Fennell, and J. A. Johnston, "The deubiquitinating enzyme USP17 is associated with non-small cell lung cancer (NSCLC) recurrence and metastasis," *Oncotarget*, vol. 4, no. 10, pp. 1836–1843, 2013.

[47] S. Ramakrishna, B. Suresh, E.-J. Lee, H.-J. Lee, W.-S. Ahn, and K.-H. Baek, "Lys-63-specific Deubiquitination of SDS3 by USP17 Regulates HDAC Activity," *Journal of Biological Chemistry*, vol. 286, no. 12, pp. 10505–10514, 2011.

[48] S. Ramakrishna, B. Suresh, and K.-H. Baek, "Biological functions of hyaluronan and cytokine-inducible deubiquitinating enzymes," *Biochimica Et Biophysica Acta-Reviews on Cancer*, vol. 1855, no. 1, pp. 83–91, 2015.

[49] S. Ramakrishna, B. Suresh, S.-M. Bae, W.-S. Ahn, K.-H. Lim, and K.-H. Baek, "Hyaluronan Binding Motifs of USP17 and SDS3 Exhibit Anti-Tumor Activity," *Plos One*, vol. 7, no. 5, 2012.

[50] B. BaculoGold, "Baculovirus Expression Vector System Manual," 1999.

[51] S. C. Gill and P. H. von Hippel, "Calculation of protein extinction coefficients from amino acid sequence data," *Analytical Biochemistry*, vol. 182, no. 2, pp. 319–326, 1989.

[52] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. Wilkins, R. Appel, and A. Bairoch, "Protein Identification and Analysis Tools on the ExpASy Server," *John M. Walker*, no. 2, pp. 571–607, 2005.

[53] S. Tayyab, S. Qamar, and M. Islam, "Size Exclusion Chromatography and Size Exclusion HPLC of Proteins," *Biochemical Education*, vol. 19, no. 3, pp. 149–152, 1991.

[54] M. P. A. Luna-Vargas, E. Christodoulou, A. Alfieri, W. J. van Dijk, M. Stadnik, R. G. Hibbert, D. D. Sahtoe, M. Clerici, V. De Marco, D. Littler, P. H. N. Celie, T. K. Sixma, and A. Perrakis, "Enabling high-throughput ligation-independent cloning and protein expression for the family of ubiquitin specific proteases," *Journal of Structural Biology*, vol. 175, no. 2, pp. 113–119, 2011.

[55] W. P. Bozza, Q. Liang, P. Gong, and Z. Zhuang, "Transient Kinetic Analysis of USP2-Catalyzed Deubiquitination Reveals a Conformational Rearrangement in the K48-Linked Diubiquitin Substrate," *Biochemistry*, vol. 51, no. 50, pp. 10075–10086, 2012.

[56] M. Li, D. Chen, A. Shiloh, J. Luo, A. Y. Nikolaev, J. Qin, and W. Gu, "Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization," *Nature*, vol. 416, no. 6881, pp. 648–653, 2002.

[57] N. D. Marchenko, S. Wolff, S. Erster, K. Becker, and U. M. Moll, "Monoubiquitylation promotes mitochondrial p53 translocation," *The EMBO Journal*, vol. 26, no. 4, pp. 923–934, 2007.

[58] B. A. Reddy, J. vanderKnaap, A. M. Bot, A. Mohd-Sarip, D. W. Dekkers, M. Timmermans, J. M. Martens, J. A. Demmers, and C. P. Verrijzer, "Nucleotide Biosynthetic Enzyme GMP Synthase Is a TRIM21-Controlled Relay of p53 Stabilization," *Molecular Cell*, vol. 53, no. 3, pp. 458–470, 2014.

[59] Z.-M. Zhang, S. Rothbart, D. Allison, Q. Cai, J. Harrison, L. Li, Y. Wang, B. Strahl, G. Wang, and J. Song, "An Allosteric Interaction Links USP7 to Deubiquitination and Chromatin Targeting of UHRF1," *Cell Reports*, vol. 12, no. 9, pp. 1400–1406, 2015.

[60] F. Colland, E. Formstecher, X. Jacq, C. Reverdy, C. Planquette, S. Conrath, V. Trouplin, J. Bianchi, V. N. Aushev, J. Camonis, A. Calabrese, C. Borg-Capra, W. Sippl, V. Collura, G. Boissy, J.-C. Rain, P. Guedat, R. Delansorne, and L. Daviet, "Small-molecule inhibitor of USP7/HAUSP ubiquitin protease stabilizes and activates p53 in cells," *Molecular Cancer Therapeutics*, vol. 8, no. 8, pp. 2286–2295, 2009.

[61] D. S. Leggett, J. Hanna, A. Borodovsky, B. Crosas, M. Schmidt, R. T. Baker, T. Walz, H. Ploegh, and D. Finley, "Multiple Associated Proteins Regulate Proteasome Structure and Function," *Molecular Cell*, vol. 10, no. 3, pp. 495–507, 2002.

[62] K. Ratia, K. S. Saikatendu, B. D. Santarsiero, N. Barretto, S. C. Baker, R. C. Stevens, and A. D. Mesecar, "Severe acute respiratory syndrome coronavirus papain-like protease: Structure of a viral deubiquitinating enzyme," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 15, pp. 5717–5722, 2006.

[63] R. Q. Kim, W. J. van Dijk, and T. K. Sixma, "Structure of USP7 catalytic domain and three Ubl-domains reveals a connector α-helix with regulatory role," *Journal of Structural Biology*, vol. 195, no. 1, pp. 11–18, 2016.

[64] L. Rougé, T. Bainbridge, M. Kwok, R. Tong, P. DiLello, I. Wertz, T. Maurer, J. Ernst, and J. Murray, "Molecular understanding of {USP7} substrate recognition and c-terminal activation," *Structure*, vol. 24, no. 8, pp. 1335 – 1345, 2016.

[65] I. H. Segel, *Enzyme Kinetics Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. New York: Wiley-Interscience Publication, 1975.

[66] Zhang, Chung, and Oldenburg, "A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.," *Journal of biomolecular screening*, vol. 4, no. 2, pp. 67–73, 1999.

[67] P. Wu, T. E. Nielsen, and M. H. Clausen, "Small-molecule kinase inhibitors: an analysis of FDA-approved drugs," *Drug Discovery Today*, vol. 21, no. 1, pp. 5–10, 2016.

APPENDIX

APPENDIX A.

A.1 USP17 Nucleotide Sequences

A.1.1 USP17 *Sf9* Optimized Nucleotide Sequence

GGATCCATGGAAGACGATTCTCTGTACCTCCGCGGTGAGTGGCAATT
CAATCACTTTTCCAAGCTGACCTCTTCCCGCCCTGATGCCGCTTTTGCTG
AAATCCAGCGCACCAGTCTCCCTGAAAAGAGCCCCCTCTCCTGCGAAAC
CAGAGTCGACCTGTGTGACGATCTGGCTCCCGTTGCTCGTCAACTCGCC
CCACGTGAGAAACTGCCACTCTCTAGCCGTCGTCCTGCTGCCGTCGGAG
CTGGACTGCAAAATATGGGCAACACTTGTTACGTCAACGCCTCCCTGCA
GTGCCTCACTTACACCCCACCATTGGCTAACTACATGCTGTCTCGTGAA
CACTCACAAACTTGCCACCGCCACAAGGGCTGTATGCTCTGCACTATGC
AAGCCCATATCACTCGCGCCCTCCACAACCCAGGCCACGTGATTCAGCC
TAGCCAAGCCTTGGCTGCCGGATTCCATCGCGGTAAACAGGAAGATGCT
CACGAGTTCCTGATGTTCACAGTGGACGCTATGAAGAAGGCCTGCTTGC
CAGGACACAAGCAGGTCGATCACCACTCCAAGGATACTACCCTGATCCA
CCAAATCTTCGGTGGTTACTGGCGTTCCCAGATCAAGTGCCTGCACTGT
CACGGTATTTCAGACACCTTCGACCCATACCTCGACATCGCCCTGGACA
TCCAGGCTGCCCAGAGTGTGCAGCAGGCTTTGGAACAGCTGGTTAAGCC
AGAAGAACTGAACGGCGAGAACGCCTACCATTGTGGCGTGTGTCTGCAA
AGGGCTCCCGCTTCCAAGACACTGACATTGCATACTAGCGCCAAGGTCC
TCATCCTCGTCTTGAAGCGTTTCTCTGACGTCACCGGTAACAAGATCGC
TAAGAACGTGCAGTACCCTGAGTGTTTGGACATGCAACCCTACATGTCC
CAGCCCAACACTGGACCCTTGGTTTACGTCTTGTATGCCGTCCTGGTGC
ATGCTGGTTGGTCATGTCACAACGGACACTACTTTAGCTACGTTAAAGC

CCAAGAGGGCCAGTGGTACAAGATGGACGACGCCGAGGTTACCGCCTCC
AGCATTACAAGCGTTCTCAGCCAACAGGCCTACGTGCTCTTCTATATCC
AGAAATCAGAGTGGGAAAGGCACTCTGAGAGCGTTAGTAGGGGTCGTG
AGCCTCGTGCTCTCGGCGCTGAGGACACCGACCGCAGAGCTACACAAGG
AGAACTCAAACGTGACCACCCCTGTTTGCAAGCTCCAGAACTGGATGAA
CACCTGGTTGAAAGGGCCACACAAGAGTCAACACTCGATCACTGGAAGT
TCCTCCAGGAGCAAAATAAGACAAAACCTGAGTTCAACGTGCGCAAGGT
GGAAGGCACCCTGCCCCCTGACGTGCTGGTTATCCACCAGTCCAAGTAC
AAGTGCGGAATGAAGAACCACCACCCTGAGCAGCAATCTTCCCTGTTGA
ACCTCAGCAGCTCAACACCTACTCACCAGGAGTCAATGAACACTGGTAC
ACTCGCCTCTCTGAGGGGTCGCGCCAGAAGGAGTAAGGGCAAGAATAAG
CATTCAAAGCGCGCTCTGCTGGTGTGCCAGTAAGAATTC

A.1.2 USP17 *E. coli* Optimized Nucleotide Sequence

ATGGAAGACGATAGCCTGTACCTGCGTGGCGAATGGCAGTTCAATCA
CTTTAGCAAACTGACCTCAAGCCGCCCGGATGCAGCGTTCGCGGAAATT
CAGCGTACCAGTCTGCCGGAAAAAAGCCCGCTGTCTTGCGAAACGCGCG
TCGATCTGTGTGATGACCTGGCACCGGTGGCACGTCAGCTGGCACCGCG
CGAAAAACTGCCGCTGAGCTCTCGTCGCCCGGCAGCAGTGGGTGCAGGT
CTGCAGAACATGGGCAATACCTGCTATGTTAACGCATCTCTGCAATGTC
TGACCTATACGCCGCCGCTGGCTAATTACATGCTGTCACGTGAACATTC
GCAGACCTGCCATCGCCACAAAGGCTGCATGCTGTGTACCATGCAAGCG
CACATTACGCGTGCCCTGCATAACCCGGGTCACGTGATCCAGCCGAGCC
AAGCACTGGCAGCTGGCTTTCATCGCGGTAAACAGGAAGATGCTCACGA
ATTTCTGATGTTCACCGTCGACGCGATGAAAAAAGCCTGTCTGCCGGGC
CATAAACAGGTTGATCATCACTCTAAAGACACCACGCTGATTCACCAGA
TCTTTGGCGGTTATTGGCGTAGCCAAATTAAATGCCTGCATTGTCACGG
TATCTCTGATACCTTCGACCCGTACCTGGATATTGCGCTGGACATCCAG

GCGGCCCAAAGCGTGCAGCAAGCCCTGGAACAGCTGGTTAAACCGGAAG
AACTGAACGGCGAAAATGCGTATCATTGCGGTGTGTGTCTGCAGCGTGC
ACCGGCTAGCAAAACCCTGACGCTGCACACCTCTGCAAAAGTCCTGATT
CTGGTGCTGAAACGCTTCAGTGATGTTACGGGCAACAAAATCGCTAAAA
ATGTCCAGTATCCGGAATGCCTGGATATGCAGCCGTACATGTCCCAACC
GAACACCGGTCCGCTGGTGTATGTTCTGTACGCAGTCCTGGTGCATGCT
GGCTGGAGTTGTCATAATGGTCACTATTTTTCCTACGTGAAAGCACAGG
AAGGCCAATGGTACAAAATGGATGACGCAGAAGTTACCGCTAGTTCCAT
TACGAGTGTTCTGTCCCAGCAAGCCTATGTCCTGTTCTACATCCAGAAA
AGCGAATGGGAACGTCATTCCGAATCAGTGTCGCGTGGTCGTGAACCGC
GTGCACTGGGTGCTGAAGATACCGACCGTCGCGCGACGCAGGGTGAACT
GAAACGTGATCATCCGTGCCTGCAAGCGCCGGAACTGGACGAACACCTG
GTTGAACGCGCCACCCAGGAATCAACGCTGGATCATTGGAAATTTCTGC
AGGAACAAAACAAAACCAAACCGGAATTCAATGTGCGCAAAGTTGAAGG
CACGCTGCCGCCGGATGTTCTGGTCATCCATCAGTCGAAATACAAATGT
GGTATGAAAAACCATCACCCGGAACAGCAATCATCGCTGCTGAATCTGA
GCTCTAGTACCCCGACGCACCAGGAATCAATGAATACCGGCACGCTGGC
CTCGCTGCGTGGTCGCGCCCGTCGTAGTAAAGGTAAAAATAAACATTCT
AAACGCGCACTGCTGGTCTGTCAA

A.2 USP7 Nucleotide Sequences

A.2.1 $USP7_{FL}$ *Sf9* Optimized Nucleotide Sequence (residues 1–1102)

AATCACCAACAGCAGCAGCAACAACAAAAAGCCGGAGAGCAACAGTT
GAGTGAACCTGAGGACATGGAGATGGAAGCAGGAGACACCGATGATCC
TCCGCGTATCACTCAAAATCCTGTGATAAACGGTAACGTCGCGCTTTCT
GATGGTCACAACACAGCGGAAGAAGACATGGAAGATGACACTTCCTGGA
GAAGCGAGGCCACCTTCCAGTTCACAGTTGAACGTTTCTCCCGCTTGAG
TGAATCCGTCCTTTCCCCTCCCTGCTTCGTGCGTAACTTGCCATGGAAA

ATCATGGTCATGCCCCGTTTCTACCCGGACAGGCCTCACCAAAAATCTG
TGGGATTCTTCCTTCAGTGCAATGCCGAGAGTGATAGCACATCTTGGTC
ATGTCACGCGCAGGCCGTGCTGAAAATTATCAACTACAGGGACGACGAG
AAGTCATTCTCCCGCCGCATCAGCCATCTTTTTTTTCATAAGGAAAACG
ATTGGGGCTTCTCAAACTTTATGGCCTGGTCCGAAGTAACCGATCCCGA
GAAAGGATTCATAGATGATGACAAAGTGACCTTCGAGGTCTTTGTTCAG
GCCGACGCTCCCCATGGCGTTGCATGGGACTCAAAGAAGCACACTGGTT
ATGTAGGCCTGAAGAACCAGGGCGCCACATGTTACATGAATAGTCTGCT
GCAGACTCTGTTCTTCACAAATCAACTGCGTAAAGCTGTTTACATGATG
CCCACTGAGGGAGATGATTCGTCGAAATCTGTTCCCCTCGCGCTTCAGC
GTGTGTTTTACGAGCTGCAGCACTCCGACAAGCCTGTGGGCACCAAGAA
ACTGACCAAGTCTTTTGGCTGGGAGACGCTTGATTCCTTCATGCAACAT
GACGTGCAAGAGCTGTGTAGGGTTCTCCTTGACAACGTCGAGAACAAGA
TGAAAGGTACCTGTGTGGAGGGCACAATCCCTAAGCTTTTTAGGGGCAA
GATGGTGTCTTATATCCAGTGCAAGGAAGTGGACTACCGTAGCGACAGG
AGGGAAGACTACTATGACATCCAGCTGTCCATCAAGGGAAAGAAGAATA
TCTTCGAATCGTTCGTAGATTATGTAGCTGTAGAACAGCTTGACGGTGA
CAACAAATACGACGCCGGTGAGCATGGACTGCAGGAAGCTGAGAAAGGT
GTGAAATTTTTGACTCTGCCGCCCGTGCTGCATTTGCAATTGATGCGCT
TCATGTACGACCCTCAAACGGATCAGAACATCAAGATCAATGACAGATT
TGAATTTCCCGAGCAGCTCCCATTGGATGAGTTCCTGCAAAAGACGGAC
CCGAAAGACCCAGCGAATTACATACTGCACGCTGTGCTGGTACACAGCG
GTGATAACCACGGTGGACACTACGTGGTATACCTCAACCCTAAAGGCGA
TGGCAAATGGTGTAAGTTCGACGATGATGTCGTTTCCCGCTGCACCAAG
GAGGAAGCGATTGAGCACAACTATGGCGGACATGACGACGACTTGTCAG
TACGCCACTGCACTAATGCGTATATGCTTGTCTATATCAGGGAAAGTAA
GTTGAGCGAGGTCCTGCAAGCTGTCACAGATCACGACATCCCGCAGCAA
TTGGTGGAGAGACTCCAAGAGGAAAAACGTATCGAGGCCCAGAAGCGTA
AGGAGAGGCAGGAAGCACATCTCTATATGCAAGTACAAATCGTTGCCGA

GGACCAGTTCTGCGGTCATCAAGGCAACGACATGTATGACGAAGAAAAG
GTTAAATACACCGTGTTTAAGGTGCTGAAGAATAGCTCCCTGGCGGAGT
TTGTACAGTCACTCAGTCAGACAATGGGTTTTCCCCAGGATCAGATCCG
CCTGTGGCCCATGCAGGCTAGATCAAATGGTACAAAACGCCCAGCTATG
CTCGATAATGAAGCCGACGGTAATAAGACCATGATCGAGTTGTCCGATA
ACGAAAATCCTTGGACAATCTTCTTGGAAACCGTGGACCCGGAGTTGGC
CGCTTCGGGCGCTACGCTGCCCAAATTCGATAAAGACCACGATGTCATG
CTGTTCCTCAAGATGTATGATCCCAAGACTAGGTCATTGAACTACTGTG
GTCATATTTACACTCCTATTAGCTGCAAGATTAGAGACTTGCTGCCTGT
TATGTGCGATCGCGCAGGCTTCATCCAGGATACCAGTCTGATCTTGTAC
GAAGAAGTGAAGCCGAACCTGACAGAACGTATCCAAGACTACGATGTGT
CACTGGACAAGGCTTTGGACGAACTCATGGATGGCGACATAATCGTCTT
CCAAAAGGACGATCCCGAAAACGACAATTCTGAGCTGCCCACTGCTAAG
GAATACTTCAGGGACTTGTACCACCGTGTCGACGTAATTTTCTGCGACA
AGACGATCCCTAACGATCCAGGTTTCGTGGTCACGCTGTCTAACAGAAT
GAACTATTTCCAAGTCGCTAAGACGGTTGCACAAAGACTGAACACTGAC
CCAATGCTCCTCCAGTTCTTCAAATCGCAAGGTTACCGCGATGGTCCAG
GAAACCCTCTCAGGCACAACTACGAGGGAACTCTGAGAGATCTCCTTCA
GTTTTTCAAGCCGAGGCAGCCCAAGAAGCTGTACTACCAACAATTGAAG
ATGAAGATTACCGACTTCGAAAACCGCAGGTCGTTTAAGTGTATCTGGT
TGAACTCTCAGTTTAGAGAAGAGGAGATCACCCTCTACCCAGACAAACA
TGGATGCGTTAGAGACCTGCTGGAAGAGTGCAAGAAGGCAGTTGAGTT
GGGAGAGAAGGCCTCCGGTAAGCTGCGCCTCCTGGAGATAGTCAGCTAC
AAGATAATTGGAGTACACCAAGAGGACGAGCTCCTCGAGTGCTTGAGCC
CAGCAACCTCACGCACCTTCAGGATTGAGGAAATCCCACTCGACCAGGT
CGATATTGATAAAGAAAATGAAATGCTCGTGACGGTTGCTCACTTCCAC
AAGGAGGTCTTCGGAACCTTCGGTATTCCTTTCCTCCTCCGTATCCATC
AGGGCGAACACTTCCGCGAGGTCATGAAACGTATTCAGAGTCTTCTCGA
CATTCAGGAGAAGGAGTTCGAGAAGTTCAAATTCGCTATCGTGATGATG

GGCAGACATCAATACATAAACGAGGACGAATACGAGGTCAACCTTAAGG
ACTTCGAGCCACAGCCAGGAAACATGTCTCACCCACGTCCATGGTTGGG
TCTTGACCACTTCAACAAGGCTCCTAAGCGCTCGCGTTACACTTATCTC
GAAAAGGCTATTAAGATTCACAACTAA

A.2.2 $USP7_{CD}$ Nucleotide Sequence (residues 208–560)

catcatcatcatcatcatgaaaacctgtattttcagggcAAAAAACATCTTCGAAAGCTTCG
TGGACTATGTGGCCGTGGAGCAGCTGGACGGCGACAACAAATACGACGC
GGGCGAACATGGTCTGCAAGAAGCCGAAAAAGGCGTTAAATTTCTGACG
CTGCCGCCGGTTCTGCACCTGCAGCTGATGCGCTTTATGTATGATCCGC
AGACCGATCAGAATATCAAAATTAACGACCGCTTTGAATTTCCGGAACA
GCTGCCGCTGGATGAATTTCTGCAAAAAACCGATCCGAAAGACCCGGCG
AACTACATTCTGCACGCGGTGCTGGTGCACAGCGGCGATAATCACGGTG
GCCACTATGTTGTTTACCTGAACCCGAAAGGCGATGGCAAATGGTGCAA
ATTTGACGATGACGTGGTTAGCCGTTGCACGAAAGAAGAAGCGATTGAA
CACAATTATGGCGGCCACGACGATGATCTGAGCGTTCGTCACTGCACCA
ATGCGTATATGCTGGTTTATATCCGCGAATCCAAACTGAGCGAAGTTCT
GCAGGCCGTTACCGACCACGATATTCCGCAGCAACTGGTTGAACGCCTG
CAGGAAGAAAAACGCATTGAAGCGCAGAAACGCAAAGAACGCCAGGAAG
CCCACCTGTATtaaggatcc

A.2.3 $USP7_{CD}$–H4–5 Nucleotide Sequence (residues 209–1102 Δ567–879)

CATATGCACCATCATCATCATCACGAGAACCTGTACTTCCAGGGTAA
AAAGCACACCGGTTACGTCGGTCTGAAGAACCAAGGTGCAACGTGTTAC
ATGAACAGCCTGCTGCAAACTCTGTTCTTCACCAACCAGCTGCGTAAGG
CTGTTTACATGATGCCTACCGAAGGTGACGACTCTTCTAAGAGCGTCCC
TCTGGCTCTGCAACGTGTATTTTACGAACTGCAACATAGCGACAAACCT
GTCGGTACTAAGAAACTGACCAAAAGCTTCGGCTGGGAAACCCTGGACT

CTTTCATGCAGCATGACGTACAGGAACTGTGTCGTGTCCTGCTGGATAA
CGTAGAAAACAAGATGAAAGGTACCTGCGTGGAGGGTACTATCCCAAAA
CTGTTCCGTGGTAAAATGGTGTCCTACATCCAGTGCAAAGAGGTGGACT
ACCGTAGCGACCGTCGTGAAGACTACTACGACATCCAACTGTCTATCAA
GGGCAAGAAAAACATCTTCGAGTCCTTCGTCGATTACGTGGCGGTGGAG
CAGCTGGATGGCGATAACAAATACGACGCAGGTGAACATGGTCTGCAGG
AAGCAGAAAAAGGTGTGAAATTCCTGACCCTGCCACCAGTACTGCATCT
GCAACTGATGCGTTTCATGTACGACCCGCAGACTGATCAGAATATCAAG
ATCAACGACCGTTTCGAGTTCCCGGAACAGCTGCCGCTGGATGAATTCC
TGCAGAAAACTGATCCGAAAGACCCGGCTAATTATATCCTGCATGCTGT
GCTGGTACACTCTGGCGATAATCACGGTGGTCACTATGTAGTATACCTG
AACCCGAAAGGCGATGGCAAATGGTGCAAATTCGACGACGATGTGGTTT
CTCGTTGTACTAAAGAGGAGGCAATCGAGCACAACTACGGTGGTCACGA
TGACGATCTGTCTGTGCGTCACTGTACTAACGCATATATGCTGGTGTAC
ATCCGTGAAAGCAAACTGTCCGAGGTTCTGCAGGCTGTTACCGATCACG
ATATCCCGCAGCAGCTGGTAGAACGTCTGCAGGAGGAAAAACGTATTGA
AGCTCAGAAACGTAAAGAACGTCAGGAAGCCCACCTGCAGCTGAAAATG
AAAATCACGGACTTCGAAAACCGTCGTTCTTTCAAATGCATCTGGCTGA
ACTCCCAGTTCCGTGAGGAAGAGATCACGCTGTATCCGGATAAACACGG
TTGCGTTCGCGATCTGCTGGAAGAATGCAAGAAAGCCGTTGAACTGGGC
GAAAAAGCGTCTGGCAAACTGCGCCTGCTGGAAATTGTTTCTTACAAAA
TCATCGGCGTTCACCAGGAAGACGAACTGCTGGAATGCCTGAGCCCGGC
CACCTCCCGTACCTTTCGCATTGAAGAAATTCCGCTGGATCAGGTTGAT
ATTGATAAAGAAAACGAAATGCTGGTTACCGTTGCGCACTTCCACAAAG
AAGTTTTCGGCACCTTCGGCATTCCGTTTCTGCTGCGCATTCACCAGGG
CGAACACTTTCGCGAAGTTATGAAACGCATTCAGTCCCTGCTGGACATC
CAGGAAAAAGAATTTGAAAAATTCAAATTTGCGATCGTTATGATGGGCC
GCCACCAGTATATCAACGAGGACGAATATGAAGTTAACCTGAAAGACTT
TGAACCGCAGCCGGGCAACATGTCCCACCCGCGCCCGTGGCTGGGCCTG

GACCACTTTAACAAAGCGCCGAAACGCTCCCGCTATACCTATCTGGAAA
AAGCGATTAAGATTCACAATTAAGGATCC

A.2.4 $USP7_{CD}$–H1–5 Nucleotide Sequence (residues 208–1102

CATATGCATCACCACCATCATCACGAAAACCTGTACTTCCAAGGTAA
GAAGCACACGGGTTACGTAGGTCTGAAGAACCAGGGTGCTACTTGCTAC
ATGAACAGCCTGCTGCAAACGCTGTTCTTCACTAACCAGCTGCGTAAGG
CAGTTTACATGATGCCTACTGAAGGTGATGATAGCAGCAAGAGCGTACC
TCTGGCACTGCAACGTGTATTCTACGAGCTGCAACATTCTGATAAGCCT
GTCGGTACTAAGAAGCTGACGAAGTCCTTCGGTTGGGAAACTCTGGATT
CTTTCATGCAGCACGACGTACAGGAGCTGTGTCGTGTACTGCTGGATAA
CGTCGAAAACAAGATGAAGGGCACTTGCGTCGAGGGTACTATCCCTAAG
CTGTTCCGTGGTAAGATGGTCTCTTACATCCAATGCAAGGAGGTCGATT
ACCGTTCTGACCGTCGTGAAGATTACTACGACATCCAGCTGTCTATCAA
GGGCAAAAAAAAACATCTTCGAGTCCTTCGTCGACTACGTGGCCGTCGAG
CAGCTGGATGGTGATAACAAATACGACGCGGGTGAGCATGGTCTGCAAG
AAGCTGAAAAAGGTGTGAAATTCCTGACGCTGCCTCCAGTGCTGCATCT
GCAACTGATGCGTTTCATGTACGATCCACAAACCGATCAGAACATCAAA
ATCAACGACCGCTTCGAGTTCCCGGAGCAACTGCCACTGGATGAATTCC
TGCAGAAAACCGATCCAAAAGACCCAGCTAACTACATCCTGCACGCAGT
GCTGGTTCACTCTGGTGATAACCACGGTGGTCATTACGTGGTTTACCTG
AACCCAAAAGGTGATGGTAAATGGTGCAAATTCGACGATGATGTGGTGT
CTCGTTGCACTAAAGAGGAGGCAATCGAGCACAACTACGGTGGTCACGA
TGACGACCTGTCTGTGCGTCACTGTACTAACGCATACATGCTGGTGTAC
ATCCGTGAGTCTAAACTGTCCGAGGTGCTGCAAGCAGTAACTGATCACG
ATATCCCGCAGCAGCTGGTTGAACGTCTGCAGGAAGAAAAACGTATCGA
GGCTCAGAAACGTAAAGAGCGTCAGGAAGCTCATCTGTACATGCAGGTG
CAGATCGTTGCAGAAGATCAGTTCTGCGGTCATCAGGGTAACGATATGT

ACGACGAAGAAAAAGTGAAATACACCGTGTTCAAAGTGCTGAAAAACTC
CTCCCTGGCTGAATTCGTGCAGTCCCTGAGCCAGACGATGGGCTTCCCG
CAGGATCAGATCCGTCTGTGGCCGATGCAGGCTCGTTCTAATGGTACTA
AACGTCCGGCAATGCTGGATAATGAAGCTGATGGTAATAAAACCATGAT
CGAACTGAGCGACAACGAAAACCCGTGGACTATCTTCCTGGAAACCGTT
GACCCGGAACTGGCTGCCTCTGGTGCGACCCTGCCGAAATTTGATAAAG
ATCATGATGTAATGCTGTTCCTGAAAATGTACGACCCGAAAACGCGTTC
CCTGAACTACTGCGGCCATATCTACACCCCGATTAGCTGTAAAATCCGC
GACCTGCTGCCGGTTATGTGTGATCGTGCGGGCTTTATTCAGGATACCT
CTCTGATTCTGTACGAAGAAGTAAAACCGAACCTGACCGAACGTATCCA
GGATTATGACGTATCCCTGGACAAAGCGCTGGATGAACTGATGGACGGC
GACATCATCGTTTTCCAGAAAGACGACCCGGAAAACGACAACAGCGAAC
TGCCGACCGCCAAAGAATATTTCCGTGACCTGTATCACCGTGTTGACGT
TATCTTCTGCGACAAAACCATCCCGAACGACCCGGGCTTTGTTGTTACC
CTGTCTAATCGTATGAACTATTTCCAGGTTGCGAAAACCGTTGCGCAGC
GTCTGAACACCGACCCGATGCTGCTGCAGTTTTTTAAAAGCCAGGGCTA
TCGTGACGGCCCGGGCAATCCGCTGCGTCACAATTATGAAGGCACCCTG
CGTGACCTGCTGCAGTTCTTTAAACCGCGTCAGCCGAAAAAACTGTATT
ATCAGCAGCTGAAAATGAAAATTACCGACTTTGAAAACCGCCGCTCTTT
TAAATGCATTTGGCTGAACTCCCAGTTCCGCGAAGAGGAGATCACCCTG
TATCCGGACAAACACGGCTGTGTTCGCGATCTGCTGGAAGAATGTAAAA
AAGCGGTTGAACTGGGCGAAAAAGCCAGCGGCAAACTGCGCCTGCTGGA
AATTGTTTCTTATAAAATCATCGGCGTTCACCAGGAAGACGAACTGCTG
GAATGCCTGTCCCCGGCGACCTCCCGCACCTTTCGCATTGAAGAAATTC
CGCTGGACCAGGTTGACATTGACAAAGAAAATGAAATGCTGGTTACCGT
TGCGCACTTTCACAAAGAAGTATTCGGCACCTTCGGCATCCCGTTTCTG
CTGCGCATTCACCAGGGCGAACACTTTCGCGAAGTAATGAAACGCATTC
AGAGCCTGCTGGACATTCAGGAAAAAGAATTCGAAAAATTCAAATTCGC
CATCGTTATGATGGGCCGCCACCAGTATATCAACGAAGACGAATATGAA

GTTAACCTGAAAGACTTCGAACCGCAGCCGGGCAACATGTCCCACCCGC
GCCCGTGGCTGGGCCTGGACCACTTTAACAAAGCGCCGAAACGCTCCCG
CTATACCTATCTGGAAAAAGCGATTAAAATTCACAACTAAGGATCC

A.2.5 **TRAF**–USP7$_{CD}$ Nucleotide Sequence (residues 1–560)

catatgCACCACCATCATCACCATGAGAACCTGTACTTCCAGGGTATGAA
CCACCAACAGCAGCAACAACAGCAAAAGGCAGGCGAGCAGCAGCTGTCT
GAACCTGAAGATATGGAAATGGAGGCTGGTGACACTGACGATCCTCCAC
GTATCACTCAGAACCCTGTTATCAACGGTAACGTGGCTCTGTCTGATGG
TCACAACACCGCAGAGGAAGACATGGAGGATGATACTAGCTGGCGCTCT
GAAGCAACTTTCCAGTTCACTGTCGAGCGTTTCTCTCGTCTGTCCGAAT
CTGTCCTGTCTCCACCATGTTTCGTACGTAACCTGCCATGGAAGATCAT
GGTCATGCCGCGTTTCTACCCGGATCGTCCGCATCAGAAATCTGTAGGT
TTCTTTCTGCAGTGTAACGCAGAATCCGATTCCACTTCCTGGTCTTGTC
ACGCACAGGCAGTACTGAAGATTATCAACTACCGTGACGACGAGAAGTC
CTTCTCCCGTCGTATCAGCCATCTGTTCTTCCACAAGGAGAACGACTGG
GGTTTCTCTAACTTCATGGCGTGGTCTGAGGTCACTGATCCGGAAAAGG
GTTTCATCGACGACGATAAGGTAACCTTCGAGGTCTTCGTTCAGGCTGA
CGCTCCGCACGGTGTAGCTTGGGATTCTAAGAAACATACCGGTTACGTT
GGTCTGAAAAACCAGGGTGCTACGTGTTACATGAACAGCCTGCTGCAGA
CCCTGTTTTTCACCAACCAGCTGCGTAAAGCGGTATACATGATGCCGAC
CGAAGGTGATGACTCTTCTAAAAGCGTTCCGCTGGCTCTGCAGCGTGTA
TTTTACGAACTGCAGCATTCCGATAAACCGGTGGGTACGAAAAAACTGA
CCAAATCCTTCGGCTGGGAAACCCTGGACTCTTTCATGCAGCACGACGT
TCAGGAACTGTGCCGTGTACTGCTGGATAACGTTGAAAACAAAATGAAA
GGCACCTGCGTGGAAGGCACCATCCCGAAACTGTTCCGTGGTAAAATGG
TGTCCTACATCCAGTGCAAAGAAGTGGACTACCGCTCCGACCGTCGTGA
AGACTACTACGACATCCAGCTGAGCATCAAAGGCAAAAAAAACATCTTC

GAAAGCTTCGTGGACTATGTGGCCGTGGAGCAGCTGGACGGCGACAACA
AATACGACGCGGGCGAACATGGTCTGCAAGAAGCCGAAAAAGGCGTTAA
ATTTCTGACGCTGCCGCCGGTTCTGCACCTGCAGCTGATGCGCTTTATG
TATGATCCGCAGACCGATCAGAATATCAAAATTAACGACCGCTTTGAAT
TTCCGGAACAGCTGCCGCTGGATGAATTTCTGCAAAAAACCGATCCGAA
AGACCCGGCGAACTACATTCTGCACGCGGTGCTGGTGCACAGCGGCGAT
AATCACGGTGGCCACTATGTTGTTTACCTGAACCCGAAAGGCGATGGCA
AATGGTGCAAATTTGACGATGACGTGGTTAGCCGTTGCACGAAAGAAGA
AGCGATTGAACACAATTATGGCGGCCACGACGATGATCTGAGCGTTCGT
CACTGCACCAATGCGTATATGCTGGTTTATATCCGCGAATCCAAACTGA
GCGAAGTTCTGCAGGCCGTTACCGACCACGATATTCCGCAGCAACTGGT
TGAACGCCTGCAGGAAGAAAAACGCATTGAAGCGCAGAAACGCAAAGAA
CGCCAGGAAGCCCACCTGTATtaaggatcc

A.2.6 H1–5 Nucleotide Sequence (residues 560–1102)

CATATGCACCATCATCATCATCACGAAAACCTGTACTTCCAGGGTGA
AGCACATCTGTACATGCAGGTCCAGATCGTCGCAGAAGATCAGTTCTGT
GGTCATCAGGGTAACGATATGTACGACGAAGAAAAGGTCAAGTACACGG
TCTTCAAGGTGCTGAAGAACTCCAGCCTGGCAGAATTCGTGCAGTCTCT
GTCTCAGACTATGGGTTTCCCTCAGGATCAGATCCGTCTGTGGCCTATG
CAAGCACGTTCTAATGGTACTAAACGTCCAGCTATGCTGGATAACGAGG
CAGATGGTAACAAGACCATGATCGAGCTGTCTGACAACGAGAACCCATG
GACTATCTTCCTGGAGACTGTGGACCCAGAACTGGCTGCTTCTGGTGCT
ACTCTGCCAAAATTTGATAAAGACCACGACGTGATGCTGTTCCTGAAAA
TGTACGACCCGAAAACGCGTTCCCTGAACTACTGCGGTCATATCTACAC
CCCAATCTCTTGCAAAATCCGCGACCTGCTGCCGGTTATGTGTGATCGT
GCTGGTTTTATCCAGGATACTAGCCTGATCCTGTACGAAGAGGTGAAAC
CGAACCTGACTGAACGTATCCAGGATTACGACGTGTCTCTGGATAAAGC

CCTGGATGAACTGATGGACGGTGACATCATCGTGTTCCAGAAAGACGAC
CCGGAAAACGATAACTCCGAACTGCCGACCGCTAAAGAATACTTCCGTG
ATCTGTACCACCGTGTAGATGTGATCTTCTGTGACAAAACCATCCCGAA
CGATCCGGGTTTCGTTGTTACCCTGTCTAACCGTATGAACTACTTCCAG
GTTGCGAAAACCGTTGCGCAGCGTCTGAACACCGATCCGATGCTGCTGC
AATTTTTCAAAAGCCAGGGTTACCGTGATGGTCCTGGCAATCCGCTGCG
TCACAATTATGAAGGCACCCTGCGTGATCTGCTGCAATTCTTTAAACCC
GCGTCAACCGAAAAAGCTGTATTATCAGCAGCTGAAAATGAAAATCACG
GACTTCGAGAACCGCCGCTCCTTCAAGTGCATCTGGCTGAACAGCCAGT
TCCGTGAAGAGGAGATCACGCTGTATCCGGACAAACACGGCTGTGTACG
CGATCTGCTGGAAGAATGCAAGAAAGCCGTTGAACTGGGCGAAAAAGCG
TCTGGCAAACTGCGTCTGCTGGAAATTGTTAGCTACAAAATCATCGGCG
TACACCAGGAAGACGAACTGCTGGAATGCCTGTCCCCGGCGACCTCCCG
TACCTTTCGTATTGAAGAAATTCCGCTGGACCAAGTTGACATTGACAAA
GAAAACGAAATGCTGGTTACCGTAGCGCACTTTCACAAAGAGGTTTTCG
GCACCTTCGGCATTCCGTTTCTGCTGCGTATTCACCAAGGCGAACACTT
TCGCGAAGTAATGAAACGCATCCAGTCCCTGCTGGACATTCAGGAAAAA
GAATTCGAGAAATTCAAATTCGCGATCGTTATGATGGGCCGCCACCAGT
ATATTAACGAGGACGAATATGAAGTAAACCTGAAAGACTTCGAACCGCA
GCCGGGCAACATGAGCCACCCGCGCCCGTGGCTGGGCCTGGACCACTTT
AACAAAGCGCCGAAACGCAGCCGCTATACCTATCTGGAAAAAGCCATTA
AGATTCACAATTAAGGATCC

A.2.7 H4–5 Nucleotide Sequence (residues 890–1102)

CATATGCACCATCATCATCATCACGAGAACCTGTACTTCCAGGGTAA
CCGTCGTTCTTTCAAATGCATCTGGCTGAACTCCCAGTTCCGTGAGGAA
GAGATCACGCTGTATCCGGATAAACACGGTTGCGTTCGCGATCTGCTGG
AAGAATGCAAGAAAGCCGTTGAACTGGGCGAAAAAGCGTCTGGCAAACT

GCGCCTGCTGGAAATTGTTTCTTACAAAATCATCGGCGTTCACCAGGAA
GACGAACTGCTGGAATGCCTGAGCCCGGCCACCTCCCGTACCTTTCGCA
TTGAAGAAATTCCGCTGGATCAGGTTGATATTGATAAAGAAAACGAAAT
GCTGGTTACCGTTGCGCACTTCCACAAAGAAGTTTTCGGCACCTTCGGC
ATTCCGTTTCTGCTGCGCATTCACCAGGGCGAACACTTTCGCGAAGTTA
TGAAACGCATTCAGTCCCTGCTGGACATCCAGGAAAAAGAATTTGAAAA
ATTCAAATTTGCGATCGTTATGATGGGCCGCCACCAGTATATCAACGAG
GACGAATATGAAGTTAACCTGAAAGACTTTGAACCGCAGCCGGGCAACA
TGTCCCACCCGCGCCCGTGGCTGGGCCTGGACCACTTTAACAAAGCGCC
GAAACGCTCCCGCTATACCTATCTGGAAAAAGCGATTAAGATTCACAAT
TAAGGATCC

A.2.8 H1–2 Nucleotide Sequence (residues 560–775

GAAGCACATCTGTACATGCAGGTCCAGATCGTCGCAGAAGATCAGTT
CTGTGGTCATCAGGGTAACGATATGTACGACGAAGAAAAGGTCAAGTAC
ACGGTCTTCAAGGTGCTGAAGAACTCCAGCCTGGCAGAATTCGTGCAGT
CTCTGTCTCAGACTATGGGTTTCCCTCAGGATCAGATCCGTCTGTGGCC
TATGCAAGCACGTTCTAATGGTACTAAACGTCCAGCTATGCTGGATAAC
GAGGCAGATGGTAACAAGACCATGATCGAGCTGTCTGACAACGAGAACC
CATGGACTATCTTCCTGGAGACTGTGGACCCAGAACTGGCTGCTTCTGG
TGCTACTCTGCCAAAATTTGATAAAGACCACGGTGGACCCAGAACTGGC
TGCTTCTGGTGCTACTCTGCCAAAATTTGATAAAGACCACGACGTGATG
CTGTTCCTGAAAATGTACGACCCGAAAACGCGTTCCCTGAACTACTGCG
GTCATATCTACACCCCAATCTCTTGCAAAATCCGCGACCTGCTGCCGGT
TATGTGTGATCGTGCTGGTTTTATCCAGGATACTAGCCTGATCCTGTAC
GAAGAGGTGAAACCGAACCTGACTGAACGTATCCAGGATTACGACGTGT
CTCTGGATAAAGCCCTGGATGAACTGATGGACGGTGACATCATCGTGTT

CCAGAAAGACGACCCGGAAAACGATAACTCCGAACTGCCGACCGCTAAA
GAATACTTCCGTGATCTGTACCACC

A.2.9 H1–3 Nucleotide Sequence (residues 560–888

GAAGCACATCTGTACATGCAGGTCCAGATCGTCGCAGAAGATCAGTT
CTGTGGTCATCAGGGTAACGATATGTACGACGAAGAAAAGGTCAAGTAC
ACGGTCTTCAAGGTGCTGAAGAACTCCAGCCTGGCAGAATTCGTGCAGT
CTCTGTCTCAGACTATGGGTTTCCCTCAGGATCAGATCCGTCTGTGGCC
TATGCAAGCACGTTCTAATGGTACTAAACGTCCAGCTATGCTGGATAAC
GAGGCAGATGGTAACAAGACCATGATCGAGCTGTCTGACAACGAGAACC
CATGGACTATCTTCCTGGAGACTGTGGACCCAGAACTGGCTGCTTCTGG
TGCTACTCTGCCAAAATTTGATAAAGACCACGGTGGACCCAGAACTGGC
TGCTTCTGGTGCTACTCTGCCAAAATTTGATAAAGACCACGACGTGATG
CTGTTCCTGAAAATGTACGACCCGAAAACGCGTTCCCTGAACTACTGCG
GTCATATCTACACCCCAATCTCTTGCAAAATCCGCGACCTGCTGCCGGT
TATGTGTGATCGTGCTGGTTTTATCCAGGATACTAGCCTGATCCTGTAC
GAAGAGGTGAAACCGAACCTGACTGAACGTATCCAGGATTACGACGTGT
CTCTGGATAAAGCCCTGGATGAACTGATGGACGGTGACATCATCGTGTT
CCAGAAAGACGACCCGGAAAACGATAACTCCGAACTGCCGACCGCTAAA
GAATACTTCCGTGATCTGTACCACCGACGACCCGGAAAACGATAACTCC
GAACTGCCGACCGCTAAAGAATACTTCCGTGATCTGTACCACCGTGTAG
ATGTGATCTTCTGTGACAAAACCATCCCGAACGATCCGGGTTTCGTTGT
TACCCTGTCTAACCGTATGAACTACTTCCAGGTTGCGAAAACCGTTGCG
CAGCGTCTGAACACCGATCCGATGCTGCTGCAATTTTTCAAAAGCCAGG
GTTACCGTGATGGTCCTGGCAATCCGCTGCGTCACAATTATGAAGGCAC
CCTGCGTGATCTGCTGCAATTCTTTAAACCCGCGTCAACCGAAAAAGCT
GTATTATCAGCAGCTGAAAATGAAAATCACGGACTTCGAGAACCGCCGC
TCC

VITA

VITA

EDUCATION

- Ph.D., 2016, Department of Biological Sciences, Purdue University, West Lafayette, Indiana
- B.Sc., 2011, Department of Biochemistry, University of Wisconsin–Madison, Madison, Wisconsin

RESEARCH EXPERIENCE

Graduate Student, Purdue University, August 2011 – December 2016

Advisor: Professor Andrew D. Mesecar

- Independently optimized refolding and purification methodology for insoluble *E. coli* BL21–DE3 expressed USP17
- Designed and optimized High–Throughput Screening (HTS) assay for the identification of small molecule inhibitors of USP17
- Performed structure activity relationship (SAR) analysis for HTS identified small molecule inhibitors of both USP17 and USP7, individually
- Determined steady–state kinetic parameters for USP17, USP7, and USP7 truncations
- Utilized isothermal calorimentry techniques to analyze protein–protein and protein–small molecule interactions
- Utilized analytical ultracentrifugation and dynamic light scattering techniques to determine oligomeric state
- Performed mode of inhibition analysis for optimized inhibitors of USP7
- Determined mode of activation of USP7 with fixed–variable studies

PUBLICATIONS

- **Hjortland, N. M.**, Mesecar, A. D. (2016)Steady–state kinetic studies reveal that the anti–cancer target Ubiquitin Specific Protease 17 (USP17) is a highly efficient deubiquitinating enzyme *(In Manuscript)*
- **Hjortland, N. M.**, Patel, A., Mesecar, A. D. (2016) A Tethered–Rheostat Model for the Intramolecular Activation of the Anti–Cancer Target USP7/HAUSP by Its HAUSP Ubiquitin–Like (HUBL) Domains 4 and 5 Requires HUBL Domains 1–3. *(In Manuscript)*
- **Hjortland, N. M.**, Pepe, A. P., Weeder, J., Molland, K., and Mesecar, A. D. (2016) Decoupling the Intramolecular Activation of USP7 by its Ubiquitin-like Domains for the Development of Small Molecule Inhibitors. *(In Manuscript).*
- Baez-Santos, Y. M., Barraza, S. J., Wilson, M. W., Agius, M. P., Mielech, A. M., **Davis, N. M.**, Baker, S. C., Larsen, S. D., Mesecar, A. D. (2015). X–ray Structural and Biological Evaluation of a series of Potent and Highly Selective Inhibitors of Human Coronavirus Papain-like Proteases. *Journal of Medicinal Chemistry*, 57(6), 2393-2412.
- Branam, A. M., **Davis N. M.**, Moore, R. W., Schneider, A. J., Vezina, C. M., and Peterson, R. E. (2013). TCDD Inhibition of Canonical Wnt Signaling Disrupts Prostatic Bud Formation in Mouse Urogenital Sinus. *Toxicological Sciences*, 133(1). 42-53.

POSTER PRESENTATIONS and TALKS

- **Hjortland, N.M.**, Patel, A., and Mesecar, A. (October 2015) "Decoupling the Allosteric Activation of Ubiquitin Specific Protease 7 by its Ubiquitin–like Domains for the Development of Small Molecule Inhibitors" American Society for Biochemistry and Molecular Biology 2016 Conference. San Diego, CA.
- **Hjortland, N.M.** and Mesecar, A. (October 2015) "Targeted Inhibition of Ubiquitin Specific Proteases for Cancer Treatment." Purdue University Center for Cancer Research Scientific Retreat. West Lafayette, IN.

- **Hjortland, N.M.**. and Mesecar,A. (September 2015) "Targeted Inhibition of Ubiquitin Specific Proteases for Cancer Treatment." Midwest Enzyme Chemistry Conference. Illinois Institute of Technology. Chicago, Illinois.
- **Hjortland, N.M.**, Molland, K., Savinov, S., Mesecar, A. (November 2014) "Targeted Inhibition of Ubiquitin Specific Proteases for Cancer Treatment." Purdue University Center for Cancer Research Scientific Retreat. West Lafayette, IN.
- **Hjortland, N.M.**, Mesecar, A. (November 2013) "From *E.coli* to Baculovirus, the Expression and Purification of USP17." Purdue University Center for Cancer Research Scientific Retreat. West Lafayette, IN.
- **Hjortland, N.M.**, Mesecar, A. (October 2013) "From *E.coli* to Baculovirus, the Expression and Purification of USP17." Midwest Enzyme Chemistry Conference. Loyola University. Chicago, Illinois.
- **Hjortland, N.M.**, Mesecar, A. (October 2013) "From *E.coli* to Baculovirus, the Expression and Purification of USP17." Walther Cancer Foundation Symposium. Notre Dame, IN.
- **Hjortland, N.M.**, Mesecar, A. (April 2013) "From *E.coli* to Baculovirus, the Expression and Purification of USP17." Ubiquitin Processes and Their Role in Cancer Mini-Symposium. West Lafayette, IN.
- **Hjortland, N.M.**, Baez–Santos, Y., Mesecar, A. (April 2012) "Protease Analysis of Inhibitor Interaction with Residue Q270." Purdue University Life Sciences Spring Retreat. West Lafayette, IN.
- **Hjortland, N.M.**, Shang, S., Xiong, M. (Spring 2010) "Delivery of Anti-Cancer Agent SAHA through Liposome Encapsulation." CALS Research Symposium. Madison, WI.

FELLOWSHIPS AND AWARDS

- American Society for Biochemistry and Molecular Biology Travel Award, 2016
- Purdue Center for Cancer Research SIRG Assistantship, 2015-2016
- Purdue Center for Cancer Research Travel Award, 2016
- Purdue University Life Sciences Travel Award, 2016
- National SMART Grant, 2010-2011
- Phi Eta Sigma Honors Society Induction, 2008
- Bryon L. Walter Family Scholarship, 2007
- Rotary Club Scholarship, 2007
- Ole Nelson Falk Scholarship, 2007